MOLECULAR DESIGN, CONSTRUCTION, AND CHARACTERIZATION OF A XYLANOSOME: A PROTEIN NANOSTRUCTURE FOR BIOMASS UTILIZATION

A Dissertation
Presented to
The Academic Faculty

by

Shara D. McClendon

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in Chemical Engineering in the
School of Chemical & Biomolecular Engineering

Georgia Institute of Technology
May 2011

MOLECULAR DESIGN, CONSTRUCTION, AND CHARACTERIZATION OF A XYLANOSOME: A PROTEIN NANOSTRUCTURE FOR BIOMASS UTILIZATION

Approved by:

Dr. Rachel Chen, Advisor
School of Chemical and Biomolecular Engineering
Georgia Institute of Technology

Dr. Sheldon May
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Andreas Bommarius
School of Chemical and Biomolecular Engineering
Georgia Institute of Technology

Dr. Carson Meredith
School of Chemical and Biomolecular Engineering
Georgia Institute of Technology

Dr. Julie Champion
School of Chemical and Biomolecular Engineering
Georgia Institute of Technology

Date Approved: January 26, 2011

To my nephew and niece, Justin and Sydney
for the world is theirs for the taking

PREFACE

The research presented in this dissertation was written either with intent to publish or already published in a scientific journal. The previously published manuscript was modified to meet formatting guidelines. This dissertation features an introduction on lignocellulosic biomass, its pretreatment and hydrolysis, and a summary of cellulosomes (Chapter 1), the experimental results of this research project with discussion (Chapters 2-5), and conclusions of this dissertation along with recommendations on directions for future studies (Chapter 6).

ACKNOWLEDGEMENTS

I would first like to thank my Lord and Savior, Jesus Christ, for salvation and providing me with all I need to do His will. May every area of my life continue to glorify Him as I strive to grow closer to Him.

Many people have significantly contributed to my success at Georgia Tech. This process has been one of the most humbling and character-building experiences for me. I am deeply indebted to my advisor, Dr. Rachel Chen, who has been a source of motivation and inspiration. Her critical insight and continued support were vital during my tenure at Tech. I would also like to thank my committee members: Dr. Andy Bommarius, Dr. Carson Meredith, Dr. Sheldon May, and Dr. Julie Champion for their time, suggestions, and enthusiasm regarding my dissertation work. I would also like to thank others who have contributed to my thesis work. Dr. Zichao Mao, a former postdoc in the Chen Research Group, aided in the development of the pQSP2 plasmid for formation of the two-unit xylanosomes. Dr. Hyun-Dong Shin, a current postdoc within our lab, was a continual source of knowledge in molecular biology, protein expression, and biomass hydrolysis. I would also like to thank current and former members of the Chen research group for being such great colleagues, especially Dr. Anne Ruffing, my past officemate during my time at Tech.

Special thanks go to Dr. Gary May, Dr. Lakeshia Taite, and the FACES (Facilitating Academic Careers in Engineering and Sciences) faculty for providing me with sponsorship and mentorship in my pursuit of a doctoral degree. Their dedication and support of minority graduate students at Tech is greatly appreciated. Also, Dr. Keith

Oden was essential in expanding my professional network and cheering me on every step of the way. Enormous thanks goes to my academic family, affectionately known as the Black Graduate Student Association (BGSA). They were always a source of encouragement and were integral in the development of my leadership skills within the academic realm. I would like to single out the following people from BGSA: Ashley Johnson, Adaora Okwo, Christianna Taylor, Dr. Tammy McCoy, Lonnie and Andrea Parker, Mela Johnson, Keith Reed, and Dimitri Hughes. Thank you for all for being a shoulder to lean on, lending your ear, making me laugh, and just plain being there during this process. I pray that our time together at Tech is only the beginning of significant, life-long friendships.

I want to especially thank my family for all of their encouragement and being there for me every moment of the way. Your unwavering love strengthened my faith in God. Your support helped me to believe in myself. To my Moma, Carol, thank you for all of your prayers. I cannot articulate everything that you have done for me and I am so grateful that you are my mother. To my Dad, Charles, thank you for reminding me that I can do anything, if I just believe. Your encouragement means the world to me. To my sister, Shelia, your resilience has been a source of hope for me. Thank you for never being more than a phone call away. To my "younger" brother, Chad, thank you for always being honest with me, reminding me to not take things so seriously, and making me laugh. I would also like to thank Veronica Gravely, Lauren Bradford, Sean Grant, and the Boyd family – your support, love, and embrace of my calling has allowed me to continue in my journey. There is no mistake that you were put in my path for this very reason.

Last, but not least, I want to thank my wonderful boyfriend, Douglas A. Brooks for the conversations that kept me sane, making me smile, and loving me as I am. Our inside jokes, song writing, adventures in cooking, and Bible studies have brought so much joy into my life. I love you dearly and am looking forward to what the future holds for us.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AA	amino acid
ABF	alpha-arabinofuranosidase/beta-xylosidase
ABFDoc	alpha-arabinofuranosidase/beta-xylosidase with dockerin domain
AFEX	ammonia fiber expansion
amp	ampicillin
BMCC	bacterial microcrystalline cellulose
CBD/CBM	carbohydrate-binding domain/module
CBM2	family 2 carbohydrate-binding module
CBM35	family 35 carbohydrate-binding module
CMC	carboxymethylcellulose
Da	dalton (molecular weight)
DCB	destarched corn bran
DNA	deoxyribonucleic acid
F1BDoc	ferulic acid esterase/acetylxylan esterase with dockerin domain
FAE	ferulic acid esterase
FAEDoc	ferulic acid esterase with dockerin domain
GH	glycosyl hydrolase
G-X-S-X-G	glycine-X-serine-X-glycine
HPLC	high-performance liquid chromatography
IMAC	immobilized metal affinity chromatography
IPTG	isopropyl β-D-thiogalactopyranoside
K_A	affinity constant
K_D	dissociation constant

K_M	Michaelis-Menten constant
k_{on}	rate of association
k_{off}	rate of dissociation
LB	luria broth
PBST	phosphate-buffered saline with tween
PCR	polymerase chain reaction
PEG	polyethylene glycol
RI	refractive index
RRS	released reducing sugars
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SP	signal peptide
SP2	two-cohesin scaffolding protein
SP3	three-cohesin scaffolding protein
SPR	surface plasmon resonance
SSF	simultaneous saccharification and fermentation
V_{max}	maximum enzymatic reaction rate
WAX	wheat arabinoxylan
X10/Xyl10	glycosyl family 10 hydrolase
X11/Xyl11	glycosyl family 11 hydrolase
X10Doc	family 10 glycosyl hydrolase with dockerin domain
X11Doc	family 11 glycosyl hydrolase with dockerin domain

LIST OF GENE AND ENZYME NOMENCLATURE

abf62A	alpha-L-arabinofuranosidase
cbpA/CbpA	scaffolding protein from *Clostridium cellulovorans*
CelA	family 9 glycosyl hydrolase endoglucanse
CelS	family 48 glycosyl hydrolase/exoglucanase
cipA/CipA	scaffolding protein from *Clostridium thermocellum*
cipC/CipC	scaffolding protein from *Clostridium cellulolyticum*
ebg98	endo-beta-galactosidase
EstA/*fee1A*/Fee1A	ferulic acid esterase
fee1B/Fee1B	ferulic acid esterase/acetylxylan esterase
lpp	lipoprotein
ompA	outer membrane protein
OlpA/B/C	cell-surface scaffolding proteins
Orf2	cell-surface scaffolding protein
SdbA	cell-surface scaffolding protein
Xyl11A	family 11 glycosyl hydrolase/xylanase
XynA	xylanase/acetylxylan esterase with dockerin domain
XynC	xylanase with dockerin domain
XynY/Z	xylanase/ferulic acid esterase with dockerin domain

SUMMARY

Lignocellulosic biomass is an abundant renewable resource targeted for biofuel production. Cellulose and hemicellulose from biomass both contain fermentable sugars and other moieties that can be converted to biofuels or other commodity chemicals. Enzymatic hydrolysis of these biopolymers is a critical step in the liberation of sugars for fermentation into desired products. In nature, anaerobic microbes produce protein nanostructures called cellulosomes that efficiently degrade cellulose substrates by combining multiple enzyme activities onto a scaffolding protein. However, current enzyme cocktails used in industry contain secretomes of aerobic microbes and are not efficient enough to be highly economical. Furthermore, most bio-processes focus on cellulose, rendering hemicellulose under-utilized. The three main objectives of this dissertation are to 1) develop multi-functional, self-assembling protein nanostructures for hemicellulose degradation using the architecture provided by cellulosomes, 2) understand the self-assembly mechanism at conditions for consolidated bioprocessing applications, and 3) compare the effectiveness of structured to non-structured hemicellulases in the hydrolysis of biomass.

Xylan is a major type of hemicellulose in biomass feedstocks targeted for biofuel production. Six different xylanosomes were designed for hydrolysis of xylan within multiple biomass substrates using the cohesin-dockerin domain systems from *Clostridium thermocellum*, *Clostridium cellulovorans*, and *Clostridium cellulolyticum*. Each two-unit structure contained a xylanase for internal cleavage of the xylan backbone and one side-chain acting enzyme, either a ferulic acid esterase or bi-functional

arabinofuranosidase/xylosidase. Expansion to three-unit xylanosomes included a family 10 or 11 xylanase, a bi-functional arabinofuranosidase/xylosidase, and bi-functional ferulic acid esterase/acetylxylan esterase. These multi-functional biocatalysts were used to degrade hemicellulose-rich wheat arabinoxylan and cellulose-containing destarched corn bran. Synergistic release of soluble sugars and ferulic acid was observed with select xylanosomes and in some cases required addition of an endoglucanase and cellobiohydrolase for enhanced hydrolysis. Furthermore, a putative ferulic acid esterase gene from the soil bacterium *Cellvibrio japonicus* was characterized and its role in xylan hydrolysis investigated.

Information for the development of stable and functional cellulosome-like biocatalysts in metabolically-engineered microbes was collected using surface plasmon resonance. The protein-protein interaction of cohesin and dockerin domains for xylanosome self-assembly was examined at various temperatures and in the presence of ethanol to mimic different hydrolysis and fermentation processes and found to retain high affinities at the selected conditions. Moreover, the high-affinity interaction of cohesin and dockerin domains in the presence of non-specific proteins eliminated the need for protein purification for xylanosome construction. In addition to development of the first cellulosome-like biocatalysts targeted for hemicellulose degradation, this dissertation provides insight on possible improvements for the enzymatic hydrolysis of biomass, as well as the applicability of xylanosomes in consolidated bioprocessing.

CHAPTER 1

INTRODUCTION

1.1 Lignocellulosic Biomass

Biomass is broadly described as biological material from living or dead organisms. In industrial terms, biomass is further defined as a renewable resource commonly used for energy, fuel, or commodity chemical production. In 2007, 6.8 quadrillion BTU of United States energy consumption was obtained from renewable resources, with 53% coming from biomass [1]. With the government-sanctioned objectives outlined in the Energy Act of 2005, lignocellulosic biomass feedstocks have been specifically targeted for increased biofuels production to meet the goal of replacing 20% of petroleum-based fuel usage by 2030. Lignocellulosic biomass contains the most abundant renewable resources available on Earth, cellulose and hemicellulose, making them ideal for use in industry, while reducing the subsequent industrial environmental impact and carbon footprint.

1.1.1 Sources of lignocellulosic biomass

Lignocellulose materials are available from various sources, such as agricultural crops and residues, energy crops, deciduous and coniferous woods, pulp and paper industry waste, and biogenic municipal solid wastes [2]. In regards to biomass for biofuels, corn crops and sugar cane are currently used to produce bioethanol in the U.S. and Brazil, respectively; soybean oil is the main feedstock for biodiesel production. Table 1.1 lists different types of biomass from multiple environmental and industrial sources.

Table 1.1 Examples of lignocellulosic biomass available from various sources

Sources	**Biomass**
Agricultural Crops	Corn, sugar cane, soybean
Agricultural Residues	Corn stover, corn fiber, wheat straw, rice hulls, sugar cane bagasse
Energy Crops	Switchgrass, Bermudagrass
Industrial Waste	Mill sludges/residues, pulping liquor
Residential Waste	Municipal solid waste, paper-based trash, yard waste, construction debris
Forestry	Various tree species (Poplar, Eucalyptus, Pine), wood residues (sawdust, wood chips)

1.1.2 Composition of lignocellulosic biomass

Lignocellulosic biomass is comprised of three main biopolymers: cellulose (40 – 60%), hemicellulose (20 – 40%), and lignin (10 – 25%), with other components such as proteins, ash, and other extractives [3]. As shown in Table 1.2, the percent composition of each component varies between different sources of biomass. Composition can vary within each source itself based on variables such as growth environment and harvesting/storage conditions for crops and process variables for industrial sources. In general, approximately two-thirds of biomass dry weight is composed of carbohydrates via cellulose and hemicellulose; these sugar polymers are targeted for biofuel production, where lignin is currently utilized for power generation.

Table 1.2 Lignocellulosic biomass compositions by percent dry basis

Biomass	**Cellulose**	**Hemicellulose**	**Lignin**	**Other**[a]
Popular[1,b]	44.70	18.55	26.44	10.31
Pine[1,b]	44.55	21.90	27.67	5.87
Switchgrass[1]	31.98	25.19	18.13	5.95
Corn Stover[2]	38 – 40	28	7 – 21	4 – 23
Wheat Straw[2]	33 – 38	26 – 32	17 – 19	5 – 18
Sugarcane Bagasse[3]	45	30	12	13
Bermuda Grass[3]	25	35	6	34

[a] – extractives, proteins, and ash
[b] – Softwoods (pine) tend to yield more 6C sugars, whereas hardwoods (poplar) contain more 5C sugars. For example, pine contains 7.9% 5C and 14% 6C, whereas, poplar has 15.4% 5C and 3.2% 6C as outlined in Hamelinck, et al. (2005). [1] Reference [3], [2] Reference [4], [3] Reference [5].

Cellulose, the most abundant renewable resource, is a chain of β-1,4-glycosyl linked D-glucose monomers. Figure 1.1 shows how these water insoluble elemental fibrils interact with each other through van der Waals forces and hydrogen bonding to form cellulose microfibrils, forming both crystalline and amorphous regions within the plant cell wall. Access to these fibrils is hindered by hemicellulose, a complex and highly-substituted heterogeneous polymer containing various pentoses (D-xylose, L-arabinose), hexoses (D-glucose, D-mannose, D-galactose) and acids (D-glucuronic acid, D-galaturonic acid, acetic acid, ferulic acid, and *p*-coumaric acid). Hemicellulose frequently interacts with lignin, a non-water soluble, high molecular weight heterogeneous polymer rich in phenylpropanoid

residues (coniferyl, sinayl, and coumeryl alcohols), via cross-linkages with hydroxycinnamic residues like *p*-coumaric and ferulic acids. An example of lignin components and composition is provided in Figure 1.2. Lignin provides critical properties for biomass, such as structural support, microbial and oxidative resistance, and impermeability [6]. The interactive and polymeric nature of these components creates biomass recalcitrance, making biomass difficult to hydrolyze [7]. Evidence of this phenomenon is seen in nature – decay of wood or plant material is a process requiring days to years to complete by microorganisms.

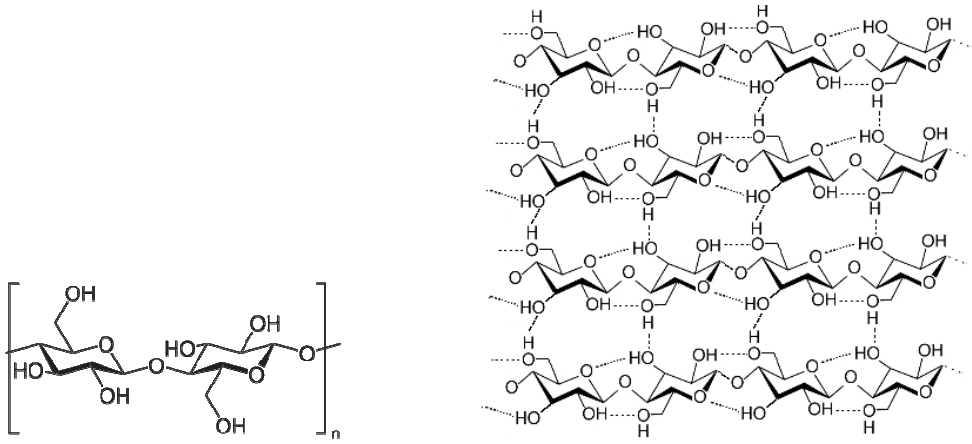


Figure 1.1 Structure of cellobiose and hydrogen boding within cellulose fibrils.

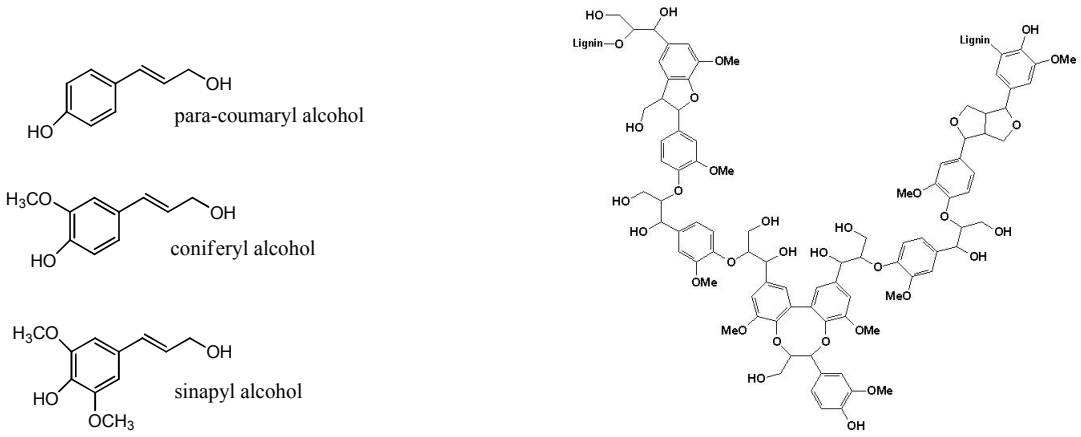


Figure 1.2 Phenylpropanoid monomers and predicted structure of lignin.

Cellulose structure is uniform throughout biomass, with variations due to different degrees of crystallinity. Lignin is not crystalline, nor has a definitive structure, but its composition is based upon the percentage of various monomer G-, S-, and H- units, where each unit is one of the monomer alcohols: **G**uaiacyl (two methoxyl groups, coniferyl alcohol), **S**yringyl (one methoxy group, sinapyl alcohol), and **H**ydroxyphenyl (no methoxy groups, *p*-coumaryl alcohol).

Hemicellulose is non-crystalline and can be present either as xylan, glucuronoxylan, arabinoxylan or glucomannan. Glucomannans and xylans are the two most common forms of hemicellulose; softwoods usually contain glucomannans, while hardwoods contain xylans [5]. Only consisting of an average of 100 monomer units per chain, hemicellulose is much smaller than cellulose and is non-crystalline. Xylans consist of β-1,4 linked D-xylopyranose units and structures vary in the nature and degree of branching of the main xylopyranosyl chain [8]. For instance, arabinoxylan has branching at the C2 and C3 positions with α-**L**-arabinofuranose, while glucuronoxylan has 4-*O*-methylglucuronic acid branching on the xylan backbone [8]. Wood xylans can exist as arabino-4-*O*-methylglucuronoxylans as seen

in softwoods or as *O*-acetyl-4-*O*-methylglucuronxylans in hardwoods. Furthermore, xylans in hardwoods have a higher degree of polymerization (150-200) than those in softwoods (70-130) [9]. In general, the xylan backbone is substituted with arabinofuranoside, glucuronosyl, and acetyl residues. Hydroxycinnamic acids, such as ferulic or *p*-coumaric acids, are ester-linked to arabinofuranosyl side chains to form disulfide bonds to cross-link xylan chains. Figure 1.3 gives examples of a few different types of hemicellulose structures.

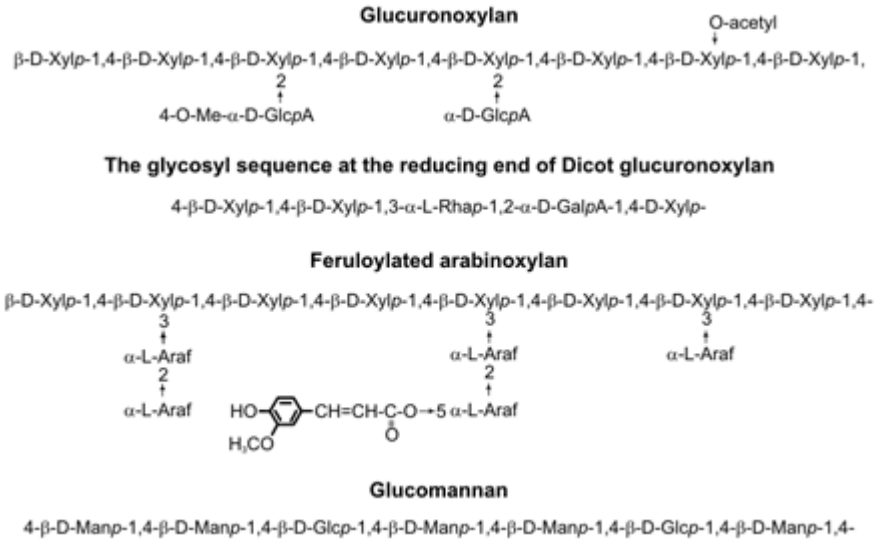


Figure 1.3 Structures of various types of hemicellulose [10].

1.1.3 Uses of lignocellulosic biomass

Solar energy captured within biomass provides a significant source of carbon that can be converted to alternative fuels. In addition to biofuels, the polymers contained in biomass are rich in sugars and acids, which can be used to synthesize high-value products and commodity chemicals which replace petroleum-based organic chemicals with bio-based

derivatives. Designing biorefineries to manufacture biofuels in conjunction with commodity products will help overcome the unfavorable economics of biofuel production [11].

1.1.3.1 Select commodity chemicals

1.1.3.1.1 Xylitol

Similar to sucrose in sweetness, xylitol is currently used as an alternative natural sweetener or food additive. In addition to sweetness, its ability to inhibit cavity-causing bacteria has encouraged xylitol usage in "sugarless" confectionaries and oral health products, such as toothpaste and mouthwash. Beyond its market in the food and health industries, xylitol can be used as a synthetic building block, making it an ideal value-added product from biomass. Multiple *E. coli* strains have been engineered to produce xylitol from mixtures of glucose and xylose present in biomass hydrolysates [12-14].

1.1.3.1.2 Lactic acid

Lactic acid is a carboxylic acid generated from pyruvate during fermentation via lactate dehydrogenase. (L)-lactic acid is the biologically relevant isomer, present in dairy foods, such as yogurt and kefir, and produced in animals during strenuous exercise. *Lactobacillus* sp. and other lactic acid bacteria are efficient at producing lactic acid [15]. Other glucose-utilizing microorganisms, such as *E. coli*, have been metabolically engineered to utilize xylose, such as that released from hemicellulose hydrolysates, to selectively yield (L)-lactic acid [16-18], (D)-lactic acid [19, 20], or a combination of both isomers [21] . Biodegradable polymers created out of lactic acid (i.e., polylactic acid) are favorable for replacing some petroleum-based polymers. Plastics made from pure or a combination of (D)-

or (L)-lactic acid can provide varying degrees of crystallinity, presenting a wide range of application in the food packaging, medical, and textile industries.

1.1.3.1.3 Ferulic acid

Hemicellulose and lignin serve as sources of ferulic acid, which can be used in various food, cosmetic, and pharmaceutical industries. Ferulic acid is a precursor in the bioproduction of vanillin, a widely used additive for food flavoring or fragrances. Furthermore, its antioxidant properties make it an interesting candidate for cosmetic application, such as sunscreen or anti-aging creams [22]. In pharma, ferulic acid isolated from biomass can be used to eliminate steps in various drug syntheses [23]. Ferulic acid was successfully removed from corn bran using the secretome of *Neosartoya spinosa* NRRL 185 for bioconversion to vanillin using *Streptomyces setonii*, giving an example of possible auxiliary processes at biorefineries in conjunction with ethanol production [24].

1.1.3.2 Biofuels

There are several biofuels being currently investigated for aiding in the replacement or supplementation of petroleum-based fuels, with biogas, ethanol, butanol, and biodiesel being the most common. While biofuels will not completely account for all transportation needs, it will aid in reducing our dependence on foreign oil.

Biogas generated from anaerobic microbial fermentations on sewage or landfill biomass generally consists of methane and carbon dioxide [25, 26]. Conversely, biogas made from the gasification of biomass is composed of hydrogen, nitrogen, and carbon monoxide [27]. Organic fractions of municipal solid waste have been also successfully converted to hydrogen [28, 29]. These gas mixtures can be used to generate electrical or

mechanical power, especially at sewage or landfill facilities for sustainability. Similar to natural gas, bio-gas provides a fuel for both heating and combustion.

Ethanol, the major focus of biofuel research, is currently blended up to 10% with petroleum-based gasoline, replacing methyl tertiary butyl ether (MTBE) as an oxygenate in the United States. Bioethanol produced in the U.S. is derived from corn crops, whereas in Brazil, a country with a sound ethanol infrastructure, sugar cane bagasse is the feedstock for bioethanol. Once biomass (cellulose) is hydrolyzed, microorganisms such as *Escherichia coli*, *Saccharomyces cerevisae,* or *Zymomonas moblis* can ferment glucose into ethanol. Recombinantly engineered species of these microbes, such as *E. coli* K011, are able to also ferment pentose sugars, such as xylose or arabinose, released during biomass hydrolysis of hemicellulose, into ethanol, increasing the overall yield of biofuel from carbon available in biomass [30-33]. Moreover, other bacteria, such as *Zymobacter palmae* and *Klebsiella oxytoca*, have also been engineered to convert xylose and/or xylo-oligosaccharides into ethanol [34, 35].

While a significant portion of research has focused on increasing ethanol yields from microbes, butanol and higher-carbon alcohols are gaining consideration for automotive fuels. Compared to ethanol, butanol is less hygroscopic and possesses a larger energy density, making it a more attractive substitute for gasoline. Solvent-tolerant microbes like *Clostridium acetobutylicum* are targeted for industrial production of butanol via Acetone-Butanol-Ethanol (ABE) fermentations, using biomass hydrolysates as carbon sources for growth [36]. Furthermore, metabolic engineering of existing amino-acid biosynthesis pathways in *E. coli* allowed synthesis of straight- and branched-chain butanols from renewable carbon sources [37].

In the case of biodiesel, plant-derived oils, such as soybean or palm oils, are trans-esterified with alcohol to give long-chain fatty acid alkyl esters. Moreover, algae have been identified as a promising source for biodiesel production, providing an advantageous method for sequestering carbon dioxide from the environment. Unlike ethanol, biodiesel can be distributed using the current fuel storage and transport systems and engines, eliminating the need for an infrastructure overhaul.

1.2 Biomass Pretreatment and Hydrolysis

In order to use lignocellulosic biomass, both naturally and industrially, its carbohydrate polymers must first be hydrolyzed into their basic components: sugars and acids. Due to the recalcitrant properties of lignocellulose, biomass hydrolysis is the most difficult and expensive aspect of biomass utilization. Thus, pretreatments are employed to reduce the complex framework of biomass and improve the enzymatic digestibility of cellulose for chemical and biofuels production. Effective pretreatments should 1) alter biomass structure, 2) remove lignin, 3) preserve hemicellulose, 4) reduce the crystalline structure of cellulose, and 5) produce no or minimal toxic by-products. In addition to these targets, these processes must minimize energy and chemical demands to be economical [38]. Three main types of pretreatments are physical, chemical, and microbial processes. While no single pretreatment meets all of these requirements, each has its own advantages and disadvantages.

1.2.1 Physical pretreatments

Generally, all biomass feedstock undergo some type of physical pretreatment, mainly for easier transport and storage, while also increasing surface area available for other processes, such as chemical pretreatment or enzymatic hydrolysis. The main types of

physical pretreatment include mechanical systems as well as processes using benign chemicals, specifically water. Mechanical treatments simply decrease biomass particle size for downstream processes and facilitate biomass transportation. Examples are ball-milling, wet-milling, and wood chipping.

In addition to reducing particle size, water treatments, such as steam explosion and liquid hot water processes, typically solubilize lignin and/or hemicellulose and enhance enzyme digestibility. Steam explosion is the rapid heating of biomass via high-pressure steam followed by a period of holding; the pressure is subsequently released for cooling and disruption of biomass structure [39]. The incubation period facilitates hydrolysis of hemicellulose, while the decompression step increases the pore size of the biomass structure [40]. Liquid hot water (LHW) treatments involve pressurized-systems to keep water liquid at elevated temperatures and can be implemented in co-current, counter-current, and flow-through systems. These pretreatments can dissolve between 40 – 60% of added biomass, including all of the hemicellulose initially present, for downstream hydrolysis and fermentation. Variability in biomass solubilization with water depends not on pretreatment conditions, but instead on the composition of biomass treated, as high levels of lignin can hinder sugar recovery, especially those found in hemicellulose [38].

During both types of treatments, water removes acidic groups present in hemicellulose, which can further contribute to the removal of polysaccharides from hemicellulose, but also degrade sugars into inhibitory aldehydes, mainly furfural and 5-hydroxymethylfurfural. Attempts to control the pH of water treatments with potassium hydroxide has proven to reduce the hydrolysis of oligosaccharides to monosaccharides thus preventing formation of aldehyde toxins [41].

1.2.2 Chemical pretreatments

Acid and alkaline pretreatments can effectively remove either hemicellulose or lignin from biomass, respectively. Chemicals used in these processes: sulfuric acid, lime, and ammonia are cheap and widely available, making these pretreatments attractive for industrial biorefineries. However, chemical recovery and recycle is of major concern, as well as post-processing steps required before enzymatic hydrolysis or microbial fermentation.

1.2.2.1 Dilute acid treatments

Acid pretreatment with dilute sulfuric acid has been shown to improve the enzyme digestibility of a wide variety of biomass substrates, including corn, rice, wheat, rye, and a number of grasses [42-47]. Other catalysts, such as hydrochloric, nitric, and phosphoric acids, have also been investigated for biomass pretreatment. A typical continuous-flow process involves subjecting a slurry of biomass (5 – 15% w/v) in 0.25 – 1.5% H_2SO_4 to temperatures from 120 – 200^{o}C for 15 – 90 minutes. Batch processes can be used for higher solids loadings of up to 40% (w/w) and lower temperatures (< 160^{o}C) [2]. Furthermore, neutralization and/or overliming of treated biomass are usually necessary prior to enzymatic saccharification or fermentation [44, 45]. Significant improvements in digestibility are observed with increasing acid concentration; however, incubation times longer than one hour tend to give minimal release of additional sugars [44, 45]. It is well known that dilute acid treatments are very effective at removing the majority of hemicellulose from the plant cell wall, exposing cellulose fibrils. Nevertheless, as with steam or LHW treatments, acids can degrade sugars into the toxins furfural and HMF that must be removed for optimal productivity of downstream processes.

1.2.2.2 Alkali pretreatments

Alkaline treatments are very efficient at catalyzing the oxidation of ether and ester linkages present within biomass, specifically lignin. Pretreatment, which is comprised of spraying an alkaline solution or slurry onto biomass, can be performed at a range of temperatures ($25^{o}C$ – $120^{o}C$) but require longer incubation times (hours – days) compared to acid pretreatments [38]. Moreover, adding oxygen to the reaction improves delignification of biomass with high lignin content [48]. Ammonium, calcium, potassium and sodium hydroxide have all been used as reactants for biomass pretreatments. However, calcium hydroxide, better known as lime, has advantages by being a cheaper, safer chemical. Furthermore, it is recoverable via reaction with carbon dioxide to form insoluble calcium carbonate and recycled to back lime by kiln techniques [49]. Lime has been successfully used in the treatment of biomass with varying levels of lignin [49, 50].

1.2.2.3 Ammonia fiber expansion (AFEX)

Ammonia fiber expansion, or AFEX, has emerged as superior method for pretreatment of herbaceous and agricultural biomass for subsequent enzymatic hydrolysis [51-55]. However, it has been found to be less effective on hardwood and softwoods [38]. This process entails combining biomass with liquid ammonia, heating the mixture and releasing the ammonia gas to “expand” the biomass structure. Also, in addition to lignin removal, AFEX is effective at decrystallizing cellulose, which is ideal prior to enzymatic hydrolysis for higher sugar recovery from biomass. Similar to other alkali treatments, AFEX leaves hemicellulose intact, making it ideal for processes utilizing both five- and six-carbon sugars. Furthermore, AFEX treatments eliminate the need for detoxification steps prior to hydrolysis and provide a nitrogen nutrient source for fermentations [53, 54].

Table 1.3 Effect of chemical pretreatments on composition and structure of lignocellulosic biomass [38]

Treatment	Increase accessible surface area	Removes hemicellulose	Removes lignin	Alters lignin structure
Steam explosion	**X**	**X**		x
Liquid hot water	**X**	**X**		x
FT liquid hot water	**X**	**X**	x	x
Dilute acid	**X**	**X**		**X**
FT dilute acid	**X**	**X**	x	**X**
AFEX	**X**	x	**X**	**X**
Alkaline/Lime	**X**	x	**X**	**X**

X – Major effect, x – minor effect, FT – flow through process

While chemical pretreatments provide cheap, fast, and effective methods of modifying biomass for hydrolysis, the resulting biomass particles or slurries must be neutralized or detoxified prior to enzymatic hydrolysis or microbial fermentation. Fungal pretreatment of biomass has emerged as a possible alternative to chemical methods.

1.2.3 Microbial (fungal) pretreatment

White-rot fungi, such as *Ceripoiopsis subvermispora* and *Phanerochaete chrysosporium*, have been researched for their ability to delignify lignocellulosic biomass [7, 56-58]. These microbes produce very efficient laccases and peroxidases necessary to disrupt the ester and ether linkages present in lignin [59]. The fungal pretreatment process starts by inoculating piles of biomass, such as wood chips, with the appropriate culture and allowing

the fungi to grow from thirty to sixty days [60]. Remaining biomass is then used for chemical or enzymatic hydrolysis and subsequent fermentation into bioproducts. In some cases, a two-step fungal pretreatment includes an initial white-rot fungi incubation followed by a brown-rot fungi, which has the ability to degrade cellulose and hemicellulose, but leave lignin intact [60]. Fungal pretreatment eliminates the need for any type of additional neutralization or detoxification steps, as with chemical methods. However, the overall time required for significant delignification makes this type of pretreatment not conducive for industrial application. Microbial pretreatments, such as those with white- and brown-rot fungi, may be best employed in smaller, regional biorefineries that have their own biomass sources and can economically produce moderate quantities of biofuels or other bio-based chemicals.

1.2.4 Enzymatic hydrolysis

The saccharification process allows the generation of smaller oligosaccharides and monosaccharides for fermentation into ethanol or other chemicals. After pretreatment, the remaining cellulose and hemicellulose within the biomass is subjected to hydrolysis with commercially available enzyme cocktails, such as those from Novozyme and Genecor, or chemically with strong acid solutions. The enzymatic route is preferred as hydrolysates can immediately be used downstream. Nature provides a multitude of enzymes and mechanisms for degrading biomass; microbial cellulases, hemicellulases, and lignases are specifically designed by nature to target their respective substrates.

1.2.4.1 Mechanisms of enzymatic hydrolysis

Cellulolytic microorganisms typically employ one of two methods for biomass degradation: 1) release of copious amounts of cellulases and hemicellulases into the

extracellular media or 2) formation of a protein nanostructure comprised of cellulases and hemicellulases called cellulosomes. Aerobic microbes live with an abundance of energy and therefore are known to have a large consortium of extracellular enzymes. This mode is very common in fungi. Secretomes of filamentous fungi, such as *Trichoderma*, have been shown to be highly effective at biomass hydrolysis and provide the source for industrial enzyme cocktails [61-64]. With minimal energy available, some anaerobic microbes, such *Clostridium thermocellum*, strategically maximize their enzyme production via cellulosomes [65]. Another minor form of biomass degradation is found in the cellulolytic bacteria *Cytophaga hutchisonii*, which possibly combines a gliding motility with cellulose utilization; it does not possess genes encoding for processive endoglucanases, cellulosome components nor does it accumulate reducing sugars in its extracellular environment [66].

1.2.4.2. Enzymes required to hydrolyze biomass

Despite the simplicity of cellulose, three different classes of enzymes are required to completely break it down into its glucose monomer. First, endoglucanases act on internal β-1,4 glycosidic bonds within cellulose, creating terminal chain ends. Next, exoglucanases, also called cellobiohydrolases, cleave cellobiose or smaller oligosaccharides from either the reducing or non-reducing end of the cellulose chain. Lastly, β-glucosidases release glucose from cellobiose and smaller oligosaccharides that cells can use for growth and energy.

Due to its complex structure, lignin degradation is classified as non-specific, requiring both oxidative peroxidases and laccases to oxidize the phenylpropanoid residues within the plant cell wall. Peroxidases are categorized into lignin (LiP) and manganese-dependent (MnP), requiring veratryl alcohol or Mn(II) as mediators, respectively [6].

Laccases are a family of blue-copper oxidases that can oxidize phenolic but not non-phenolic lignin components [2].

Hemicellulose is more heterogeneous in nature, and thus, requires more enzymatic activities for full degradation. The group of necessary enzymes depends on the composition of hemicellulose and the side groups present on the polymer backbone. For example, xylan can contain various pentoses and hexoses, as well as acid moieties. Overall, six different types of enzymatic activities shown in Figure 1.4 are needed for xylan hydrolysis: xylanase, xylosidase, arabinofuranosidase, glucuronidase, acetylxylan esterase, and feruloyl esterase [67, 68].

A summary of all enzymatic activities needed to break down cellulose, hemicellulose, and lignin is provided in Table 1.5.

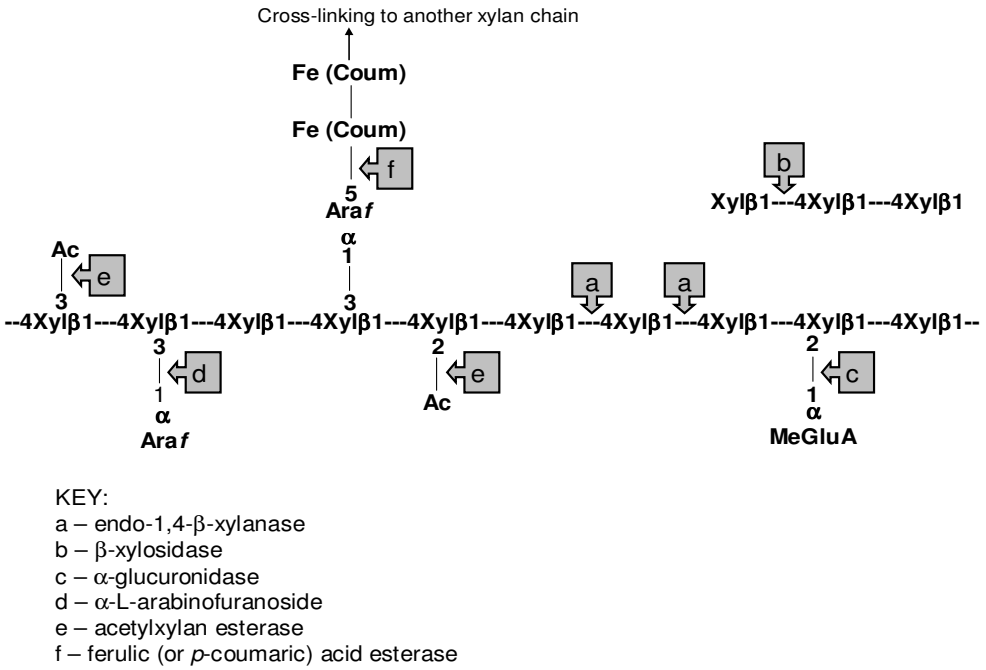


Figure 1.4 Schematic of enzymatic digestion of xylan.

Table 1.4 List of enzyme activities required for biomass degradation

Biopolymer	Enzyme	EC Number	Mode of action
Cellulose	Endoglucanase	3.2.1.4	Endo-hydrolysis of 1,4-β-D-glucosidic linkages
	Exoglucanase	3.2.1.91	Hydrolysis of 1,4-β-D-glucosidic linkages releasing cellobiose
	β-glucosidase	3.2.1.21	Hydrolyzes cellobiose and short chain oligosaccharides to glucose
Hemicellulose (Xylan)	Xylanase	3.2.1.8	Hydrolyzes β-1,4 xylose linkages within xylan backbone
	β-xylosidase	3.2.1.37	Hydrolyzes short chain xylo-oligosaccharides to xylose
	Acetylxylan esterase	3.1.1.6	Hydrolyzes ester-bonded acetyl groups present in xylans
	Arabinofuranosidase	3.2.1.55	Hydrolyzes terminal non-reducing arabinofuranose from arabinoxylans
	α-Glucuronidase	3.2.1.131	Hydrolyzes glucuronic acid from glucuronoxylans
	Ferulic acid esterase	3.1.1.73	Hydrolyzes ester-bonded feruloyl groups in xylans
Lignin	Lignin peroxidase	1.11.1.7	Oxidation of benzilic alcohols, cleavage of C-C and C-O bonds
	Manganese peroxidase	1.11.1.7	Dependent on hydrogen peroxide and Mn(II) ions
	Laccase	1.10.3.2	Oxidizes phenolic subunits of lignin

1.2.4.3 Synergism of biomass hydrolysis

The structure of biomass requires the synergistic action of lignocellulolytic enzymes, making both unstructured and cellulosome enzyme systems effective at hydrolyzing biomass.

Extracellular fractions of cellulolytic bacteria and fungi contain a variety of cellulases and hemicellulases, working together to facilitate release of monosaccharides, acids, and other phenolic compounds. Synergy between cellulases is well-known. The products of endoglucanases are the substrates for cellobiohydrolases and the products of cellobiohydrolases are the substrates for glucosidases; indeed, synergy between the various cellulases is required for effectiveness. Furthermore, combining cellulosomal cellulases and xylanases of *Clostridium cellulovorans* has shown significant improvement in the release of reducing sugars from corn residues [69]. In regards to hemicellulose, synergy between hemicellulases has also been identified. For example, research done at Novozymes led to identifying optimal cocktails of fungal xylanases, β-xylosidases, and arabinofuranosidases to effectively release arabinose and xylose from wheat arabinoxylan [64]. Similarly, combining fungal xylanases and ferulic acid esterases has also improved the release of ferulic acid from wheat bran and arabinoxylans [70, 71].

1.3 Cellulosomes

In the 1980s, Edward Bayer and Raphael Lamed first discovered the extracellular organization of cellulolytic enzymes of the thermophillic anaerobic bacterium, *Clostridium thermocellum*, in structures termed cellulosomes [72, 73]. Since then, a number of cellulolytic anaerobic bacteria and fungi are found to produce these novel protein nanostructures to maximize effectiveness of synthesized proteins in their energy-limited environment. By placing hydrolases and esterases into a scaffold, these microbes are extremely efficient at using lignocellulosic biomass as a carbon source. Evidence of cellulosome expression was initially determined by electron microscopy [74, 75], and recently using genomics [76] and proteomics [77]. Anaerobic bacteria and fungi producing

cellulosomes have been isolated from soil, rumen, sewage, compost, and wood fermenters [78]. Cellulosomes can range from around 600 kDa for single complexes to greater than 3 MDa for polycellulosomes [65]. The structural organization and main components of bacterial cellulosomes are well-understood and give insight into why these structures allow microbes to be efficient at hydrolyzing biomass.

Table 1.5 Cellulosome-producing anaerobic bacteria and fungi

Bacteria		**Fungi**
Clostridium thermocellum[c]	*Acetivibrio cellulolyticus*	*Neocalimastix* sp.
Clostridium cellulovorans	*Bacteroides cellulosolvens*	*Piromyces* sp.
Clostridium cellulolyticum[c]	*Butyrivbrio fibrisolvens*	*Orpinomyces* sp.
Clostridium josui	*Ruminococcus flavefaciens**	
Clostridium acetobutylicum[c,d]	*Ruminococcus albus*	
Clostridium papyrosolvens	*Ruminococcus succinogenes*	
Closridium cellbioparum		

[c] Genome sequence is available.
[d] The genome of *C. acetobutylicum* contains putative cellulosome genes, but this organism does not secrete cellulosomes [79].

1.3.1 Architecture of cellulosomes

Cellulosomes are self-assembling, multi-functional protein nanostructures that conduct efficient hydrolysis of cellulose. Specific cellulosome structures vary between organisms, but still possess the same basic components of scaffolding proteins and dockerin-tagged enzymes. Furthermore, some cellulosomes are attached to the cell wall, while others are free in the supernatant [80]. Compared to bacterial cellulosomes, fungal dockerin domain structures have been studied and found to vary significantly from the amino acid sequence of

their bacteria counterparts [81]. However, a scaffolding protein has yet to be identified from a cellulosome-producing fungus, and as a result, not much is known concerning the overall structure of fungal cellulosomes [65]. Therefore, this section will focus on the architecture of cellulosomes from bacteria.

Figure 1.5 shows a basic schematic of the cellulosome from *C. thermocellum*, the most extensively studied structure out of all cellulosome-producing bacteria. The main feature of the cellulosome architecture is the self-assembly mechanism, provided by the interaction of type I cohesin domains within scaffolding proteins (CipA) and type I dockerin domains within enzymes. In addition, cellulosome cell surface attachment (SdbA, OlpB, Orf2) and possible polycellulosome formation mechanisms (OlpB, Orf2, Cthe_0736) are allowed via a separate and distinct type II cohesin and dockerin domain interaction. Interestingly, single enzymes can also be displayed on the cell surface via OlpA and OlpC. Thus, the variety of scaffolding proteins and multiple types of cohesins and dockerins allow *C. thermocellum* to have dynamic and multi-functional cellulosomes for degrading recalcitrant lignocellulose.

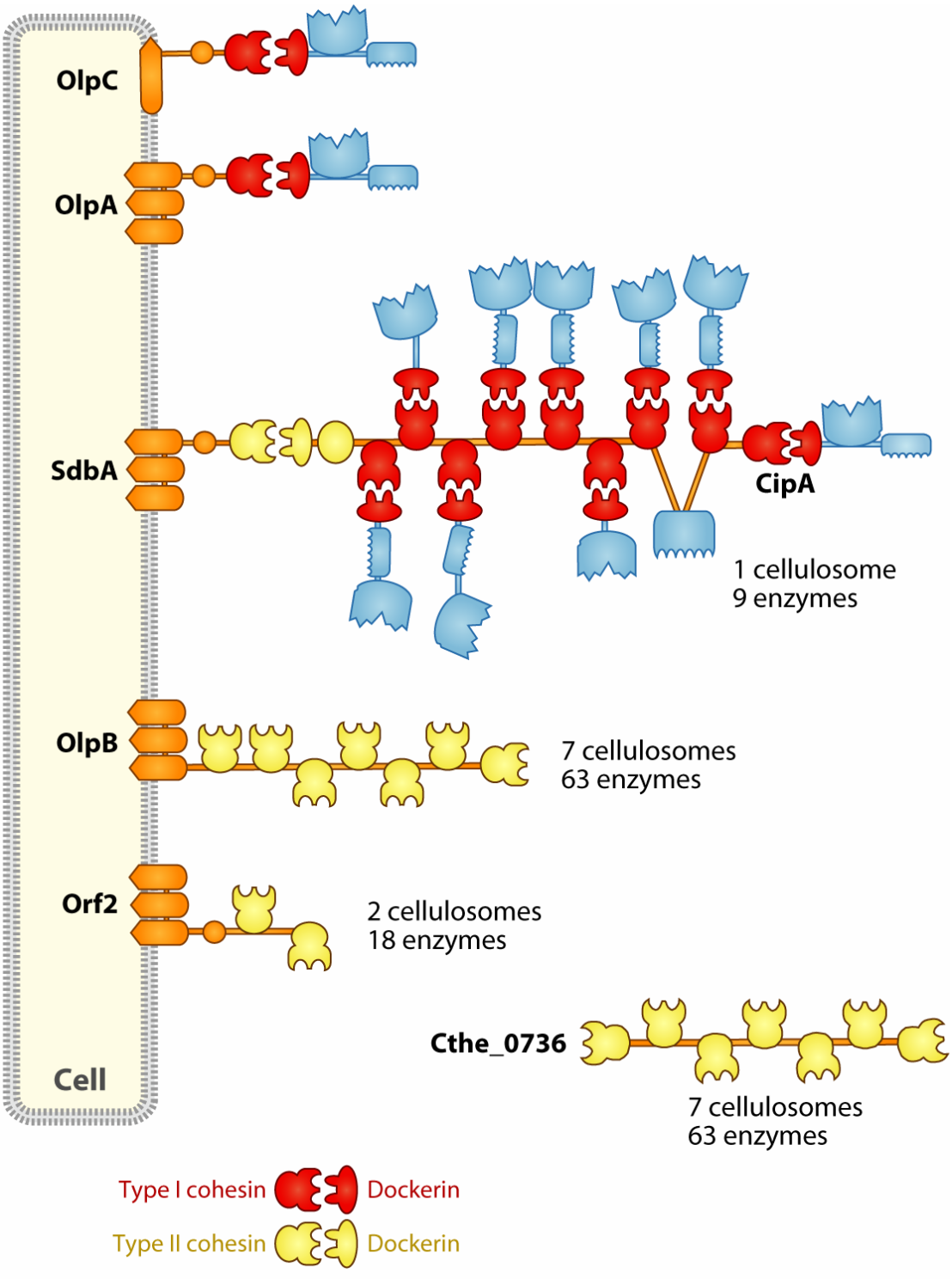


Figure 1.5 Architecture of *Clostridium thermocellum* cellulosomes [65].

1.3.1.1 Cohesin-dockerin interaction

As previously mentioned, cellulosomes assemble and attach to the cell surface via cohesin and dockerin domains. Cohesins are classified by amino acid sequence, while dockerins are typically grouped based on the type of cohesin with which it interacts [82]. The framework for cellulosomes is provided by scaffolding proteins, which contain multiple type I cohesin units and, if associated with the cell wall, typically type II cohesins. Scaffolding proteins can contain a variable number of cohesin units; CipA of *C. thermocellum* has nine type I cohesins, while *C. josui* CipA has only six [83]. While the cellulosome from *C. thermocellum* is the most well-known, cellulosomes of other bacteria, such as *Acetivibrio cellulolyticus* [84], *Bacteroides cellulosolvens* [85], and *Ruminococcus flavefaciens* [86], show significant divergence in scaffolding protein composition, including the addition of a type III cohesin/dockerin system.

Cohesin and dockerin domains interact in a high-affinity, calcium-dependent, protein-protein interaction ($K_A \sim 10^7 - 10^{10}$ M^{-1}) allowing cellulosomal components to be secreted and resulting cellulosomes constructed in the extracellular environment [87]. Type I dockerins attached to enzymes interact with corresponding cohesins to form cellulosomes, while type II dockerins interact with scaffolding proteins to link cellulosomes to the cell wall. Cohesins are generally around 140 – 150 amino acid residues, forming a "jelly-roll" nine-stranded β-sandwich. Dockerin domains are between 60 – 70 amino acids and contain two highly-conserved duplicate segments of about 22 residues with similarity to the EF-hand motif. Crystal structures of a type I cohesin and dockerin from *C. thermocellum* revealed that two F-hand motif helices of the dockerin domain carry Ca^{2+} ions, while the cohesin domain has no calcium binding sites [88, 89]. Interestingly, in the absence of calcium ions,

the dockerin domain loses stability and undergoes a conformational change from the F-hand domain structure, resulting in the calcium-dependent binding mechanism between dockerins and cohesins [90]. The cohesin is believed to interact with the dockerin domain through a network of hydrogen bonding between a conserved aspartate residue on one of the cohesin β-strands, serine and threonine residues on a dockerin α-helix, and water molecules [91].

The presence of two α-helices within dockerins, one from each repeated 22-amino acid sequence, allows for a dual binding mode in cellulosome assembly. Site-directed mutagenesis studies revealed that alteration of the either helix did not eliminate the ability of dockerin domains to interact with cohesins, conferring plasticity in cellulosome formation [92, 93]. With multiple binding locations, this feature is suggested to help facilitate reorganization of enzymes within cellulosomes to eliminate spatial hindrances between enzymes or changes in available substrate composition.

Another useful feature of cellulosome assembly is the species-specificity of the cohesin-dockerin interaction. This interaction has been observed through a number of techniques, including non-denaturing PAGE, affinity blotting, surface plasmon resonance and isothermal titration calorimetry. A thorough investigation of both type I and type II cohesin-dockerin interactions from multiple cellulosome-producing species revealed that type I domains show almost exclusive specificity, while type II domains allow extensive cross-species interactions [82]. However, cross-species type I interactions between *C. thermocellum* and *C. cellulolyticum* or *C. josui* have been observed, although these interactions are much weaker than intra-species interactions [94, 95]. Furthermore, changing a threonine to a leucine removes species-specificity of a *C. thermocellum* dockerin to a *C. cellulolyticum* cohesin, revealing the sensitivity of this feature [96].

1.3.1.2 Carbohydrate-binding modules

Neither cohesins within scaffolding proteins nor dockerins possess any type of catalytic activity or interact with the target biomass substrate. However, carbohydrate-binding domains or modules (CBD, CBM) provide another facet to cellulosomes by attaching to various carbohydrates, allowing the entire complex to be optimally positioned for efficient hydrolysis. Due to the variety of polysaccharides and crystalline structures within lignocellulosic biomass, a range of CBMs are available within the cellulosome. Family 3a CBMs found within the scaffolding proteins of *C. thermocellum* (CipA) and *C. cellulovorans* (CbpA) bind specifically to crystalline cellulose. These type A CBMs have flat binding sites that are ideal for interacting with the planar surface of crystalline cellulose [97]. CBMs are also found in the various modular cellulosomal enzymes for proper substrate positioning for enzymatic activity. In addition to family 3a, CBMs from families 3b, 3c, and 22 present within cellulosomal enzymes bind to amorphous cellulose, smaller oligosaccharides, and xylan, respectively [98, 99]. Type B CBMs (family 3b, 22) are known as 'chain binders' because of their ability to interact with oligosaccharides with a degree of polymerization between 3 and 6. Additionally, type C CBMs (family 3c) prefer mono-, di- or tri-saccharides [97].

1.3.1.3 Cellulosomal enzymes

Since lignocellulosic biomass is a very complex substrate, cellulosomes contain a vast array of enzymatic artillery for its hydrolysis. Quantitative proteomic analyses have identified the majority of cellulosomes are comprised of endoglucanases and cellobiohydrolases when grown on cellulose substrates [100, 101]. Hemicellulases are also incorporated, commonly with at least two enzymatic activities, such as XynY/Z from *C.*

thermocellum which contains both xylanase and ferulic acid esterase catalytic domains [102] and a bi-functional xylanase/acetylxylan esterase from *C. cellulovorans* [103]. Other enzymatic activities such as mannases, chitinases, and pectate lyases have also been found in *C. thermocellum* cellulosomes showing the diverse consortia of enzymes needed during hydrolysis of lignocellulose [104-106].

Depending on the carbon source, cellulosomal compositions change to accommodate differences in biomass complexity. For instance, when grown on cellulose, cellulosomes of *C. thermocellum* include mainly endoglucanases and exoglucanases from glycosyl hydrolyase families 5, 8, 9 and 48 with CelS (GH48) and CelA (GH 9) being the most abundant enzymes [100, 101]. However, when grown on cellobiose, the percentage of hemicellulases increases from 12% to 22%, indicative of a type of "scavenger" mode where cellulosomes are used to expose cellulose buried under hemicellulose [101]. Thus, it has been suggested that a form of substrate-specific regulation or catabolite repression of cellulosomal components exists, helping the microbe to adapt to the available growth source [100, 101, 105].

1.3.2 Designer cellulosomes

Significant work has been done on designing and constructing miniature forms of cellulosomes. These designer cellulosomes are less than one-third the size of native cellulosomes, but have shed light on the overall mechanism of efficient hydrolysis of cellulose. Using recombinant DNA techniques, chimeric scaffolding proteins and dockerin-tagged enzymes have been cloned and expressed into *E. coli*, purified, and used for cohesin-dockerin analysis and cellulose degradation. Taking advantage for species-specificity, cohesins and dockerin-tagged enzymes from *C. thermocellum*, *C. cellulolyticum*, and *R.*

flavefaciens were used to construct bi- and tri-functional designer cellulosomes. These structures gave superior hydrolysis of insoluble substrates, such as Avicel and BMCC, and behaved similar to free enzyme systems on more soluble bacterial cellulose and phosphoric-acid swollen cellulose [107]. Furthermore, bi-functional cellulosomes were used to determine optimal synergy between family 9 and family 48 glycosyl hydrolases, as well as the substrate targeting effects of CBMs in hydrolysis of insoluble cellulose. Overall, designer cellulosomes are inferior to native cellulosomes in performance, exhibiting between 3- and 10-fold lower activity on bacterial cellulose and Avicel, respectively [107, 108]. However, increasing enzymatic diversity can improve this gap, as tri-functional cellulosomes incorporating a xylanase were 3-fold more active on hatched straw than those containing only endo- and exo-glucanases [108].

Non-cellulosomal enzymes, such as a family 6 cellulase from *Neocallimastix patriciarum* and xylanases from *Thermobifida fusca*, have successfully been integrated into designer cellulosomes [109, 110]. Moreover, conversion of non-cellulosomal enzymes from *C. thermocellum* to cellulosomal components provide potential for expanding the repertoire of enzymes available for incorporation into chimeric cellulosomes [111].

Most studies perform construction of designer cellulosomes *in vitro*, but there are a few reports that suggest a trend moving toward a more streamlined approach. Recombinant *Bacillus subtilis* strains secreting either a single cohesin scaffolding protein, an endoglucanase, or a xylanase from *C. cellulovorans* were co-cultured to produce mini-cellulosomes *in vivo* [112]. Similarly, *C. acetobutylicum* was engineered to express one- and two-unit heterologous cellulosomes [113, 114]. Cell-surface display of tri-functional scaffolding proteins on *Saccharomyces cerevisiae* has shown simultaneous saccharification

and fermentation of phosphoric-acid swollen cellulose into ethanol. In these cases, chimeric cellulosomal enzymes were either co-expressed with scaffolding protein [115] or added separately after scaffolding protein expression [116].

Designer cellulosomes have the potential to improve enzymatic hydrolysis of biomass for biofuel and bio-based chemical production. Further work needs to be done in generating mini-cellulosome systems that can match or surpass degradation of biomass with native cellulosomes. In order to meet this target, expansion of mini-cellulosomes to include more enzymes, such as hemicellulases and perhaps ligninases, will be required, especially in hydrolysis of industrially-relevant biomass feedstocks of bio-energy crops and agricultural residues. Furthermore, using the framework provided by cellulosomes and mini-cellulosomes, protein nanostructures targeted for other substrates besides cellulose would be beneficial in further advancement of the enzymatic hydrolysis of biomass. Since hemicellulose contains fermentable sugars, development of a nanostructure specifically designed for its hydrolysis could allow for increased utilization of biomass and make biorefineries more economically viable.

1.4 Project Objectives

The work described in this dissertation focuses on three objectives: 1) design and construction of a multi-functional, self-assembling protein nanostructure targeted to degrade hemicellulose, termed a xylanosome, 2) to characterize the self-assembly aspect of the xylanosome, and 3) to compare hydrolysis of lignocellulose using structured xylanosomes and corresponding free enzyme systems.

1.4.1 Design and construction of two- and three-unit xylanosomes

Hemicellulose is substrate rich in five-carbon sugars and other acids that could be of use in the production of biofuels or other value-added chemicals. However, this biopolymer is not of major focus in the utilization of biomass for biofuels. Due to its varied composition, hemicellulose requires many enzymes for complete hydrolysis. Placement of those enzymes in a scaffolding protein may enhance synergy. Therefore, protein nanostructures similar to mini-cellulosomes combining multiple hemicellulases could improve hydrolysis of hemicellulose, resulting in higher production of bio-based products from biomass.

Mini-cellulosomes have been successfully designed and constructed using various genes encoding type I cohesins, type I dockerins, and enzymes for cellulose hydrolysis [107, 117]. In addition to those used in designer cellulosomes, genes from the sequenced genomes of multiple cellulosome-producing bacteria provide a diverse pool from which to design xylanosomes. Xylan is a major type of hemicellulose and is thus selected for the target substrate of xylanosomes. Since xylan is comprised of a xylose-polymer backbone with various side chain substituents, the initial design of xylanosomes is to include two enzymes: a backbone-degrading enzyme and a side-chain cleaving enzyme. Each enzyme will possess a type I dockerin domain from a different species, which will interact with a chimeric scaffolding protein containing two corresponding divergent type I cohesin domains. For three-unit xylanosomes, a third type I cohesin and dockerin domain pair from another species will be used to expand the scaffolding protein and functionalize a different hemicellulase, respectively, and expand the capability of the xylanosome to enhance xylan hydrolysis. The species-specificity of the cohesin and dockerin domains will allow the two- and three-unit

xylanosomes to form in a 1:1 ratio of enzyme to cohesin, creating a biocatalyst with known composition.

1.4.2 Characterization of self-assembly via cohesin-dockerin systems

The high affinity protein-protein interaction between cohesins and dockerins has been investigated, confirming the species-specificity of type I cohesin and dockerin binding. With further development of designer cellulosomes and xylanosomes, especially in engineering of recombinant cellulosome-producing microorganisms, it is imperative to understand the cohesin/dockerin interaction under various stresses. Using surface plasmon resonance, I will analyze the affinity constants for selected cohesin and dockerin pairs to gain knowledge on how these nanostructures withstand industrial process conditions.

1.4.3 Lignocellulose hydrolysis using free enzymes and xylanosomes

As observed with mini-cellulosomes, increased synergy between enzymes in structured form versus the same enzymes in free form may be observed with xylan hydrolysis. The performance of constructed xylanosomes will be compared to corresponding free enzyme systems on the release of sugars and other side-chain moieties that could serve as precursors for bio-based chemicals, such as those described in Section 1.1.3. Hydrolysis of various types of biomass will be included, as well as synergies with other biomass-degrading enzymes. This work will showcase the first design, construction, and application of a chimeric protein nanostructure dedicated for xylan hydrolysis. Furthermore, it will provide insight into future directions in the improvement of enzymatic hydrolysis of biomass.

1.5 References

1. *Renewable Energy Trends in Consumption and Electricity, 2008*. 2010 [cited 2010 January 8]; Available from: http://www.eia.gov/fuelrenewable.html.

2. Saha, B.C., et al., *Lignocellulose biodegradation*. ACS symposium series 889. 2004, Washington, DC: American Chemical Society : Distributed by Oxford University Press. xii, 400 p.

3. Hamelinck, C.N., G.v. Hooijdonk, and A.P.C. Faaij, *Ethanol from lignocellulosic biomass: techno-economic performance in short-, middle- and long-term.* Biomass and Bioenergy, 2005. **28**(4): p. 384-410.

4. Reddy, N. and Y. Yang, *Biofibers from agricultural byproducts for industrial applications.* Trends Biotechnol, 2005. **23**(1): p. 22-7.

5. Saha, B.C., *Hemicellulose bioconversion.* J Ind Microbiol Biotechnol, 2003. **30**(5): p. 279-91.

6. Perez, J., et al., *Biodegradation and biological treatments of cellulose, hemicellulose and lignin: an overview.* Int Microbiol, 2002. **5**(2): p. 53-63.

7. Vicuña, R., *Ligninolysis.* Molecular Biotechnology, 2000. **14**(2): p. 173-176.

8. Sun, R., X.F. Sun, and J. Tomkinson, *Hemicelluloses and their derivatives*, in *Hemicelluloses: Science and techology*, P. Gatenholm and M. Tenkanen, Editors. 2004, ACS Division of Cellulose and Renewable Materials: Washington, D.C. p. 2-22.

9. Subramaniyan, S. and P. Prema, *Biotechnology of microbial xylanases: enzymology, molecular biology, and application.* Crit Rev Biotechnol, 2002. **22**(1): p. 33-64.

10. *Introduction to Plant Cell Walls*. 2007 August 20, 2007 [cited December 10, 2010; Available from: http://www.ccrc.uga.edu/~mao/intro/ouline.htm.

11. Ragauskas, A.J., et al., *The path forward for biofuels and biomaterials.* Science, 2006. **311**(5760): p. 484-9.

12. Cirino, P.C., J.W. Chin, and L.O. Ingram, *Engineering Escherichia coli for xylitol production from glucose-xylose mixtures.* Biotechnol Bioeng, 2006. **95**(6): p. 1167-76.

13. Khankal, R., J.W. Chin, and P.C. Cirino, *Role of xylose transporters in xylitol production from engineered Escherichia coli.* J Biotechnol, 2008. **134**(3-4): p. 246-52.

14. Khankal, R., et al., *Comparison between Escherichia coli K-12 strains W3110 and MG1655 and wild-type E. coli B as platforms for xylitol production.* Biotechnol Lett, 2008. **30**(9): p. 1645-53.

15. Reddy, G., et al., *Amylolytic bacterial lactic acid fermentation -- A review.* Biotechnology Advances. **26**(1): p. 22-34.

16. Dien, B.S., N.N. Nichols, and R.J. Bothast, *Recombinant Escherichia coli engineered for production of L-lactic acid from hexose and pentose sugars.* J Ind Microbiol Biotechnol, 2001. **27**(4): p. 259-64.

17. Dien, B.S., N.N. Nichols, and R.J. Bothast, *Fermentation of sugar mixtures using Escherichia coli catabolite repression mutants engineered for production of L-lactic acid.* J Ind Microbiol Biotechnol, 2002. **29**(5): p. 221-7.

18. Zhou, S., K.T. Shanmugam, and L.O. Ingram, *Functional replacement of the Escherichia coli D-(-)-lactate dehydrogenase gene (ldhA) with the L-(+)-lactate dehydrogenase gene (ldhL) from Pediococcus acidilactici.* Appl Environ Microbiol, 2003. **69**(4): p. 2237-44.

19. Shukla, V.B., et al., *Production of D(-)-lactate from sucrose and molasses.* Biotechnol Lett, 2004. **26**(9): p. 689-93.

20. Zhou, S., et al., *Production of optically pure D-lactic acid in mineral salts medium by metabolically engineered Escherichia coli W3110.* Appl Environ Microbiol, 2003. **69**(1): p. 399-407.

21. Eiteman, M.A., et al., *A substrate-selective co-fermentation strategy with Escherichia coli produces lactate by simultaneously consuming xylose and glucose.* Biotechnol Bioeng, 2009. **102**(3): p. 822-7.

22. Mathew, S. and T.E. Abraham, *Ferulic acid: an antioxidant found naturally in plant cell walls and feruloyl esterases involved in its release and their applications.* Crit Rev Biotechnol, 2004. **24**(2-3): p. 59-83.

23. Ou, S. and K.-C. Kwok, *Ferulic acid: pharmaceutical functions, preparation and applications in foods*. 2004, John Wiley & Sons, Ltd. p. 1261-1269.

24. Shin, H.D., et al., *A complete enzymatic recovery of ferulic acid from corn residues with extracellular enzymes from Neosartorya spinosa NRRL185.* Biotechnol Bioeng, 2006. **95**(6): p. 1108-15.

25. Amon, T., et al., *Biogas production from maize and dairy cattle manure--Influence of biomass composition on the methane yield.* Agriculture, Ecosystems & Environment, 2007. **118**(1-4): p. 173-182.

26. Lastella, G., et al., *Anaerobic digestion of semi-solid organic waste: biogas production and its purification.* Energy Conversion and Management, 2002. **43**(1): p. 63-75.

27. Wang, T., J. Chang, and P. Lv, *Synthesis Gas Production via Biomass Catalytic Gasification with Addition of Biogas.* Energy & Fuels, 2005. **19**(2): p. 637-644.

28. Dong, L., et al., *Hydrogen production characteristics of the organic fraction of municipal solid wastes by anaerobic mixed culture fermentation.* International Journal of Hydrogen Energy, 2009. **34**(2): p. 812-820.

29. Lay, J.-J., Y.-J. Lee, and T. Noike, *Feasibility of biological hydrogen production from organic fraction of municipal solid waste.* Water Research, 1999. **33**(11): p. 2579-2586.

30. Becker, J. and E. Boles, *A modified Saccharomyces cerevisiae strain that consumes L-Arabinose and produces ethanol.* Appl Environ Microbiol, 2003. **69**(7): p. 4144-50.

31. Lawford, H.G. and J.D. Rousseau, *Performance testing of Zymomonas mobilis metabolically engineered for cofermentation of glucose, xylose, and arabinose.* Appl Biochem Biotechnol, 2002. **98-100**: p. 429-48.

32. Ohta, K., et al., *Genetic improvement of Escherichia coli for ethanol production: chromosomal integration of Zymomonas mobilis genes encoding pyruvate*

decarboxylase and alcohol dehydrogenase II. Appl Environ Microbiol, 1991. **57**(4): p. 893-900.

33. Sedlak, M. and N.W. Ho, *Production of ethanol from cellulosic biomass hydrolysates using genetically engineered Saccharomyces yeast capable of cofermenting glucose and xylose.* Appl Biochem Biotechnol, 2004. **113-116**: p. 403-16.

34. Burchhardt, G. and L.O. Ingram, *Conversion of xylan to ethanol by ethanologenic strains of Escherichia coli and Klebsiella oxytoca.* Appl Environ Microbiol, 1992. **58**(4): p. 1128-33.

35. Yanase, H., et al., *Genetic engineering of Zymobacter palmae for production of ethanol from xylose.* Appl Environ Microbiol, 2007. **73**(8): p. 2592-9.

36. Qureshi, N., et al., *Butanol Production from Corn Fiber Xylan Using Clostridium acetobutylicum*. 2006, American Chemical Society. p. 673-680.

37. Atsumi, S., T. Hanai, and J.C. Liao, *Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels.* Nature, 2008. **451**(7174): p. 86-9.

38. Mosier, N., et al., *Features of promising technologies for pretreatment of lignocellulosic biomass.* Bioresour Technol, 2005. **96**(6): p. 673-86.

39. Zimbardi, F., et al., *Lignocellulosic biomass as carbon source by steam explosion pretreatment.* New Biotechnology, 2009. **25**(Supplement 1): p. S275-S275.

40. Glasser, W.G. and R.S. Wright, *Steam-assisted biomass fractionation. II. fractionation behavior of various biomass resources.* Biomass and Bioenergy, 1998. **14**(3): p. 219-235.

41. Weil, J., et al., *Continuous pH monitoring during pretreatment of yellow poplar wood sawdust by pressure cooking in water.* Applied Biochemistry and Biotechnology, 1998. **70-72**(1): p. 99-111.

42. Dien, B.S., et al., *Chemical composition and response to dilute-acid pretreatment and enzymatic saccharification of alfalfa, reed canarygrass, and switchgrass.* Biomass and Bioenergy, 2006. **30**(10): p. 880-891.

43. Karimi, K., G. Emtiazi, and M.J. Taherzadeh, *Ethanol production from dilute-acid pretreated rice straw by simultaneous saccharification and fermentation with Mucor indicus, Rhizopus oryzae, and Saccharomyces cerevisiae.* Enzyme and Microbial Technology, 2006. **40**(1): p. 138-144.

44. Saha, B.C., et al., *Dilute acid pretreatment, enzymatic saccharification, and fermentation of rice hulls to ethanol.* Biotechnol Prog, 2005. **21**(3): p. 816-22.

45. Saha, B.C., et al., *Dilute acid pretreatment, enzymatic saccharification and fermentation of wheat straw to ethanol.* Process Biochemistry, 2005. **40**(12): p. 3693-3700.

46. Sun, Y. and J.J. Cheng, *Dilute acid pretreatment of rye straw and bermudagrass for ethanol production.* Bioresource Technology, 2005. **96**(14): p. 1599-1606.

47. Torget, R., et al., *Dilute-Acid Pretreatment of Corn Residues and Short-Rotation Woody Crops.* Applied Biochemistry and Biotechnology, 1991. **28-29**(1): p. 75-86.

48. Chang, V. and M. Holtzapple, *Fundamental factors affecting biomass enzymatic reactivity.* Applied Biochemistry and Biotechnology, 2000. **84-86**(1): p. 5-37.

49. Chang, V., M. Nagwani, and M. Holtzapple, *Lime pretreatment of crop residues bagasse and wheat straw.* Applied Biochemistry and Biotechnology, 1998. **74**(3): p. 135-159.

50. Chang, V., et al., *Oxidative lime pretreatment of high-lignin biomass.* Applied Biochemistry and Biotechnology, 2001. **94**(1): p. 1-28.

51. Chundawat, S.P., B. Venkatesh, and B.E. Dale, *Effect of particle size based separation of milled corn stover on AFEX pretreatment and enzymatic digestibility.* Biotechnol Bioeng, 2007. **96**(2): p. 219-31.

52. Kim, Y., et al., *Enzyme hydrolysis and ethanol fermentation of liquid hot water and AFEX pretreated distillers' grains at high-solids loadings.* Bioresour Technol, 2008. **99**(12): p. 5206-15.

53. Lau, M.W. and B.E. Dale, *Cellulosic ethanol production from AFEX-treated corn stover using Saccharomyces cerevisiae 424A(LNH-ST).* Proc Natl Acad Sci U S A, 2009. **106**(5): p. 1368-73.

54. Lau, M.W., B.E. Dale, and V. Balan, *Ethanolic fermentation of hydrolysates from ammonia fiber expansion (AFEX) treated corn stover and distillers grain without detoxification and external nutrient supplementation.* Biotechnol Bioeng, 2008. **99**(3): p. 529-39.

55. Teymouri, F., et al., *Ammonia fiber explosion treatment of corn stover.* Appl Biochem Biotechnol, 2004. **113-116**: p. 951-63.

56. Cullen, D., *Recent advances on the molecular genetics of ligninolytic fungi.* Journal of Biotechnology, 1997. **53**(2-3): p. 273-289.

57. Martinez, A.T., et al., *Biodegradation of lignocellulosics: microbial, chemical, and enzymatic aspects of the fungal attack of lignin.* Int Microbiol, 2005. **8**(3): p. 195-204.

58. Tanaka, H., et al., *Degradation of wood and enzyme production by Ceriporiopsis subvermispora.* Enzyme and Microbial Technology, 2009. **45**(5): p. 384-390.

59. Bonugli-Santos, R.C., et al., *Production of laccase, manganese peroxidase and lignin peroxidase by Brazilian marine-derived fungi.* Enzyme and Microbial Technology. **46**(1): p. 32-37.

60. Lee, J.-W., et al., *Enzymatic saccharification of biologically pretreated Pinus densiflora using enzymes from brown rot fungi.* Journal of Bioscience and Bioengineering, 2008. **106**(2): p. 162-167.

61. Boisset, C., et al., *Optimized mixtures of recombinant Humicola insolens cellulases for the biodegradation of crystalline cellulose.* Biotechnol Bioeng, 2001. **72**(3): p. 339-45.

62. de Vries, R.P., et al., *Synergy between enzymes from Aspergillus involved in the degradation of plant cell wall polysaccharides.* Carbohydr Res, 2000. **327**(4): p. 401-10.

63. Rosgaard, L., et al., *Evaluation of minimal Trichoderma reesei cellulase mixtures on differently pretreated Barley straw substrates.* Biotechnol Prog, 2007. **23**(6): p. 1270-6.

64. Sorensen, H.R., et al., *Enzymatic hydrolysis of wheat arabinoxylan by a recombinant "minimal" enzyme cocktail containing beta-xylosidase and novel endo-1,4-beta-*

xylanase and alpha-l-arabinofuranosidase activities. Biotechnol Prog, 2007. **23**(1): p. 100-7.

65. Fontes, C.M. and H.J. Gilbert, *Cellulosomes: highly efficient nanomachines designed to deconstruct plant cell wall complex carbohydrates.* Annu Rev Biochem, 2010. **79**: p. 655-81.

66. Xie, G., et al., *Genome sequence of the cellulolytic gliding bacterium Cytophaga hutchinsonii.* Appl Environ Microbiol, 2007. **73**(11): p. 3536-46.

67. Dodd, D. and I.K. Cann, *Enzymatic deconstruction of xylan for biofuel production.* Glob Change Biol Bioenergy, 2009. **1**(1): p. 2-17.

68. Shallom, D. and Y. Shoham, *Microbial hemicellulases.* Curr Opin Microbiol, 2003. **6**(3): p. 219-28.

69. Kosugi, A., K. Murashima, and R.H. Doi, *Characterization of two noncellulosomal subunits, ArfA and BgaA, from Clostridium cellulovorans that cooperate with the cellulosome in plant cell wall degradation.* J Bacteriol, 2002. **184**(24): p. 6859-65.

70. Faulds, C.B., et al., *Synergy between xylanases from glycoside hydrolase family 10 and family 11 and a feruloyl esterase in the release of phenolic acids from cereal arabinoxylan.* Appl Microbiol Biotechnol, 2006. **71**(5): p. 622-9.

71. Faulds, C.B. and G. Williamson, *Release of ferulic acid from wheat bran by a ferulic acid esterase (FAE-III) from Aspergillus niger.* Appl Microbiol Biotechnol, 1995. **43**(6): p. 1082-7.

72. Bayer, E.A., E. Setter, and R. Lamed, *Organization and distribution of the cellulosome in Clostridium thermocellum.* J Bacteriol, 1985. **163**(2): p. 552-9.

73. Lamed, R., E. Setter, and E.A. Bayer, *Characterization of a cellulose-binding, cellulase-containing complex in Clostridium thermocellum.* J Bacteriol, 1983. **156**(2): p. 828-36.

74. Mayer, F., et al., *Macromolecular Organization of the Cellulolytic Enzyme Complex of Clostridium thermocellum as Revealed by Electron Microscopy.* Appl Environ Microbiol, 1987. **53**(12): p. 2785-92.

75. Bayer, E.A. and R. Lamed, *Ultrastructure of the cell surface cellulosome of Clostridium thermocellum and its interaction with cellulose.* J Bacteriol, 1986. **167**(3): p. 828-36.

76. Zverlov, V.V., J. Kellermann, and W.H. Schwarz, *Functional subgenomics of Clostridium thermocellum cellulosomal genes: identification of the major catalytic components in the extracellular complex and detection of three new enzymes.* Proteomics, 2005. **5**(14): p. 3646-53.

77. Cho, W., et al., *Cellulosomic profiling produced by Clostridium cellulovorans during growth on different carbon sources explored by the cohesin marker.* J Biotechnol. **145**(3): p. 233-9.

78. Doi, R.H., et al., *Cellulosomes from mesophilic bacteria.* J Bacteriol, 2003. **185**(20): p. 5907-14.

79. Sabathe, F., A. Belaich, and P. Soucaille, *Characterization of the cellulolytic complex (cellulosome) of Clostridium acetobutylicum.* FEMS Microbiol Lett, 2002. **217**(1): p. 15-22.

80. Dror, T.W., et al., *Regulation of expression of scaffoldin-related genes in Clostridium thermocellum.* J Bacteriol, 2003. **185**(17): p. 5109-16.

81. Ljungdahl, L.G., *The cellulase/hemicellulase system of the anaerobic fungus Orpinomyces PC-2 and aspects of its applied use.* Ann N Y Acad Sci, 2008. **1125**: p. 308-21.

82. Haimovitz, R., et al., *Cohesin-dockerin microarray: Diverse specificities between two complementary families of interacting protein modules.* Proteomics, 2008. **8**(5): p. 968-79.

83. Kakiuchi, M., et al., *Cloning and DNA sequencing of the genes encoding Clostridium josui scaffolding protein CipA and cellulase CelD and identification of their gene products as major components of the cellulosome.* J Bacteriol, 1998. **180**(16): p. 4303-8.

84. Xu, Q., et al., *A novel Acetivibrio cellulolyticus anchoring scaffoldin that bears divergent cohesins.* J Bacteriol, 2004. **186**(17): p. 5782-9.

85. Xu, Q., et al., *Architecture of the Bacteroides cellulosolvens cellulosome: description of a cell surface-anchoring scaffoldin and a family 48 cellulase.* J Bacteriol, 2004. **186**(4): p. 968-77.

86. Jindou, S., et al., *Conservation and divergence in cellulosome architecture between two strains of Ruminococcus flavefaciens.* J Bacteriol, 2006. **188**(22): p. 7971-6.

87. Fierobe, H.P., et al., *Cellulosome from Clostridium cellulolyticum: molecular study of the Dockerin/Cohesin interaction.* Biochemistry, 1999. **38**(39): p. 12822-32.

88. Lytle, B.L., et al., *Solution structure of a type I dockerin domain, a novel prokaryotic, extracellular calcium-binding domain.* J Mol Biol, 2001. **307**(3): p. 745-53.

89. Tavares, G.A., P. Beguin, and P.M. Alzari, *The crystal structure of a type I cohesin domain at 1.7 A resolution.* J Mol Biol, 1997. **273**(3): p. 701-13.

90. Michiels, J., et al., *The functions of Ca(2+) in bacteria: a role for EF-hand proteins?* Trends Microbiol, 2002. **10**(2): p. 87-93.

91. Handelsman, T., et al., *Cohesin-dockerin interaction in cellulosome assembly: a single Asp-to-Asn mutation disrupts high-affinity cohesin-dockerin binding.* FEBS Lett, 2004. **572**(1-3): p. 195-200.

92. Carvalho, A.L., et al., *Evidence for a dual binding mode of dockerin modules to cohesins.* Proc Natl Acad Sci U S A, 2007. **104**(9): p. 3089-94.

93. Pinheiro, B.A., et al., *The Clostridium cellulolyticum dockerin displays a dual binding mode for its cohesin partner.* J Biol Chem, 2008. **283**(26): p. 18422-30.

94. Pages, S., et al., *Species-specificity of the cohesin-dockerin interaction between Clostridium thermocellum and Clostridium cellulolyticum: prediction of specificity determinants of the dockerin domain.* Proteins, 1997. **29**(4): p. 517-27.

95. Jindou, S., et al., *Cohesin-dockerin interactions within and between Clostridium josui and Clostridium thermocellum: binding selectivity between cognate dockerin and cohesin domains and species specificity.* J Biol Chem, 2004. **279**(11): p. 9867-74.

96. Mechaly, A., et al., *Cohesin-dockerin interaction in cellulosome assembly: a single hydroxyl group of a dockerin domain distinguishes between nonrecognition and high affinity recognition.* J Biol Chem, 2001. **276**(13): p. 9883-8.

97. Boraston, A.B., et al., *Carbohydrate-binding modules: fine-tuning polysaccharide recognition.* Biochem J, 2004. **382**(Pt 3): p. 769-81.

98. Ali, E., et al., *Functions of family-22 carbohydrate-binding module in Clostridium thermocellum Xyn10C.* Biosci Biotechnol Biochem, 2005. **69**(1): p. 160-5.

99. Schwarz, W.H., *The cellulosome and cellulose degradation by anaerobic bacteria.* Appl Microbiol Biotechnol, 2001. **56**(5-6): p. 634-49.

100. Gold, N.D. and V.J. Martin, *Global view of the Clostridium thermocellum cellulosome revealed by quantitative proteomic analysis.* J Bacteriol, 2007. **189**(19): p. 6787-95.

101. Raman, B., et al., *Impact of pretreated Switchgrass and biomass carbohydrates on Clostridium thermocellum ATCC 27405 cellulosome composition: a quantitative proteomic analysis.* PLoS One, 2009. **4**(4): p. e5271.

102. Blum, D.L., et al., *Feruloyl esterase activity of the Clostridium thermocellum cellulosome can be attributed to previously unknown domains of XynY and XynZ.* J Bacteriol, 2000. **182**(5): p. 1346-51.

103. Kosugi, A., K. Murashima, and R.H. Doi, *Xylanase and acetyl xylan esterase activities of XynA, a key subunit of the Clostridium cellulovorans cellulosome for xylan degradation.* Appl Environ Microbiol, 2002. **68**(12): p. 6399-402.

104. Zverlov, V.V., K.P. Fuchs, and W.H. Schwarz, *Chi18A, the endochitinase in the cellulosome of the thermophilic, cellulolytic bacterium Clostridium thermocellum.* Appl Environ Microbiol, 2002. **68**(6): p. 3176-9.

105. Han, S.O., et al., *Regulation of expression of cellulosomes and noncellulosomal (hemi)cellulolytic enzymes in Clostridium cellulovorans during growth on different carbon sources.* J Bacteriol, 2004. **186**(13): p. 4218-27.

106. Kurokawa, J., et al., *Sequence of the Clostridium thermocellum mannanase gene man26B and characterization of the translated product.* Biosci Biotechnol Biochem, 2001. **65**(3): p. 548-54.

107. Fierobe, H.P., et al., *Degradation of cellulose substrates by cellulosome chimeras. Substrate targeting versus proximity of enzyme components.* J Biol Chem, 2002. **277**(51): p. 49621-30.

108. Fierobe, H.P., et al., *Action of designer cellulosomes on homogeneous versus complex substrates: controlled incorporation of three distinct enzymes into a defined trifunctional scaffoldin.* J Biol Chem, 2005. **280**(16): p. 16325-34.

109. Mingardon, F., et al., *Incorporation of fungal cellulases in bacterial minicellulosomes yields viable, synergistically acting cellulolytic complexes.* Appl Environ Microbiol, 2007. **73**(12): p. 3822-32.

110. Morais, S., et al., *Contribution of a xylan-binding module to the degradation of a complex cellulosic substrate by designer cellulosomes.* Appl Environ Microbiol. **76**(12): p. 3787-96.

111. Vazana, Y., et al., *Interplay between Clostridium thermocellum family 48 and family 9 cellulases in cellulosomal versus noncellulosomal states.* Appl Environ Microbiol. **76**(10): p. 3236-43.

112. Arai, T., et al., *Synthesis of Clostridium cellulovorans minicellulosomes by intercellular complementation.* Proc Natl Acad Sci U S A, 2007. **104**(5): p. 1456-60.

113. Mingardon, F., et al., *Heterologous production, assembly, and secretion of a minicellulosome by Clostridium acetobutylicum ATCC 824.* Appl Environ Microbiol, 2005. **71**(3): p. 1215-22.

114. Perret, S., et al., *Production of heterologous and chimeric scaffoldins by Clostridium acetobutylicum ATCC 824.* J Bacteriol, 2004. **186**(1): p. 253-7.

115. Wen, F., J. Sun, and H. Zhao, *Yeast surface display of trifunctional minicellulosomes for simultaneous saccharification and fermentation of cellulose to ethanol.* Appl Environ Microbiol. **76**(4): p. 1251-60.

116. Tsai, S.L., et al., *Functional assembly of minicellulosomes on the Saccharomyces cerevisiae cell surface for cellulose hydrolysis and ethanol production.* Appl Environ Microbiol, 2009. **75**(19): p. 6087-93.

117. Fierobe, H.P., et al., *Design and production of active cellulosome chimeras. Selective incorporation of dockerin-containing enzymes into defined functional complexes.* J Biol Chem, 2001. **276**(24): p. 21257-61.

CHAPTER 2

MOLECULAR DESIGN AND CONSTRUCTION OF TWO-UNIT XYLANOSOMES

2.1 Abstract

Designer cellulosomes have been integral in understanding the mechanism of assembly and cellulose hydrolysis of native cellulosomes. However, protein nanostructures targeted for hemicellulose are yet to be reported in nature or by *in vitro* construction. Development of such a protein structure would be advantageous since microbes are being engineered to utilize five-carbon sugars from hemicellulose for biofuels and other value-added chemicals, such as xylitol. Using the architecture of native and designer cellulosomes, four different xylanosomes were designed and constructed using recombinant DNA and molecular cloning techniques. The main components include a two-cohesin scaffolding protein with *Clostridium thermocellum* and *Clostridium cellulovorans* cohesin domains, and four different hemicellulases: a glycosyl hydrolase family 10 xylanase, a glycosyl hydrolase family 11 xylanase, a ferulic acid esterase, and a bi-functional alpha-L-arabinofuranosidase/beta-D-xylosidase with either a *C. thermocellum* or *C. cellulovorans* dockerin domain added to the carboxyl-terminal of the catalytic domain. Each component was cloned into *Escherichia coli*, expressed, and purified using immobilized metal affinity chromatography. Proper construction of all four xylanosomes was confirmed using surface plasmon resonance.

2.2 Introduction

Enzymatic hydrolysis is a critical process in the conversion of lignocellulosic materials to biofuels and other chemicals. Cellulosomes allow anaerobic bacteria and fungi to efficiently degrade biomass by placing enzymes in close proximity to the substrate and each other, maximizing their synergy. Current tools in genomics and molecular cloning provide the necessary information and protocols to construct miniature forms of cellulosomes. These designer cellulosomes have been used to observe the interaction of cohesins and dockerins for self-assembly [1, 2], synergy of cellulosomal enzymes, and performance of structured enzymes on cellulose substrates compared to free enzymes [3-5]. First constructs combined one or two cellulases into chimeric scaffolding proteins comprised of single or a combination of type I cohesins from *Clostridium thermocellum* and *Clostridium cellulolyticum*, with and without the presence of a family 3a carbohydrate binding module. Cellulosomal enzymes were cloned with native and non-native type I dockerin domains [3]. Scaffolding proteins were then extended to contain three different cohesins, including a type I cohesin from the rumen bacteria, *Ruminococcus flavefaciens*, generating tri-functional mini-cellulosomes [4]. Creation of alternate geometries, such as cyclic, symmetrical, asymmetrical designs, demonstrated the possibility of generating higher-ordered nanostructures using cohesins and dockerins from multiple species [6]. In addition, mini-cellulosomes have been produced in *Clostridium acetobutylicum*, which has genes containing cohesin and dockerin domains, but does not naturally secrete cellulosome complexes [7].

Non-cellulosomal enzymes can also be incorporated into designer cellulosomes. Two fungal cellulases from *Neocallimastix partriciarum* were functionalized with bacterial

dockerins for incorporation into di- and tri-functional complexes, retaining their activity towards cellulose [5]. Moreover, homologous free cellulases from *C. thermocellum* were converted into cellulosomal enzymes and successfully inserted into scaffolding proteins [8]. Furthermore, inclusion of hemicellulases into these nanostructures has also been investigated. Two- and three-unit designer cellulosomes have incorporated combinations of cellulases with xylanases for improved hydrolysis of complex substrates, expanding their application beyond homogenous crystalline cellulose derivatives [4, 9].

The construction of designer cellulosomes has been confirmed primarily using non-denaturing or native polyacrylamide gel electrophoresis (native PAGE) [3, 5, 6]. Other methods such as protein-based microarrays and surface plasmon resonance (SPR) have also been employed in measuring and observing the cohesin-dockerin interaction using cellulosomal domains from various anaerobic bacteria [10, 11]. Native PAGE requires fairly cheap and readily available equipment to perform, whereas microarrays and SPR depend on more sophisticated chip design or availability of specific equipment.

While glucose from cellulose is the primary carbon source for cellulosome-producing microbes, metabolically-engineered microbes are capable of utilizing five-carbon sugars from hemicellulose present within biomass [12-15]. These biological advancements provide a need for improved enzymatic hydrolysis of hemicellulose for release of such sugars. While evidence of high molecular-weight, extracellular protein complexes containing only xylanase and beta-xylosidase activities has been found, cohesin and dockerin domains were absent, suggesting a non-cellulosomal type of structure [16-18]. Therefore, a cellulosome-like protein nanostructure for hemicellulose degradation has yet to be constructed or found in nature. The precedents of chimeric cellulosomes outline the logic and tools to build such a

biocatalyst. This chapter describes the molecular design and construction of four different two-unit xylanosomes, using either a glycosyl hydrolase (GH) family 10 or family 11 xylanase as a backbone-acting enzyme and a side-chain acting ferulic acid esterase or bi-functional α-L-arabinofuranosidase/β-D-xylosidase for the hydrolysis of xylan substrates. Recombinant *Escherichia coli* strains harboring genes for the dockerin-tagged enzymes and two-cohesin scaffolding protein were successfully used for expression of each of the five xylanosome components.

2.3 Results

2.3.1 Cloning and expression of components for two-unit xylanosome

Cellulosomes of anaerobic bacteria and fungi are extremely efficient at hydrolyzing biomass, specifically cellulose, into mono- and smaller oligosaccharides for cell growth. Based on the architecture of these nanostructures, many designer cellulosomes have been constructed using both native and chimeric cellulases and scaffolding proteins for cellulose degradation [3-6, 8, 9]. To construct a protein nanostructure targeted for xylan hydrolysis, named a xylanosome, four different dockerin-bearing enzymes and a single scaffolding protein with two different cohesins were designed, cloned and expressed into *E. coli*.

A scaffoldin protein, SP2, was constructed using the first cohesin domain from *CipA*, the primary scaffoldin protein of *C. thermocellum*, and the ninth cohesin domain of *CbpA*, the primary scaffoldin protein of *C. cellulovorans*, with linker sequences from both scaffolding proteins included. Figure 2.1 shows a schematic of the construction of SP2. Therefore, two enzymes, one bearing a *C. thermocellum* dockerin and the other a *C. cellulovorans* dockerin, can be used to construct a two-unit xylanosome, taking advantage of the species-specificity of the cohesin-dockerin interaction.

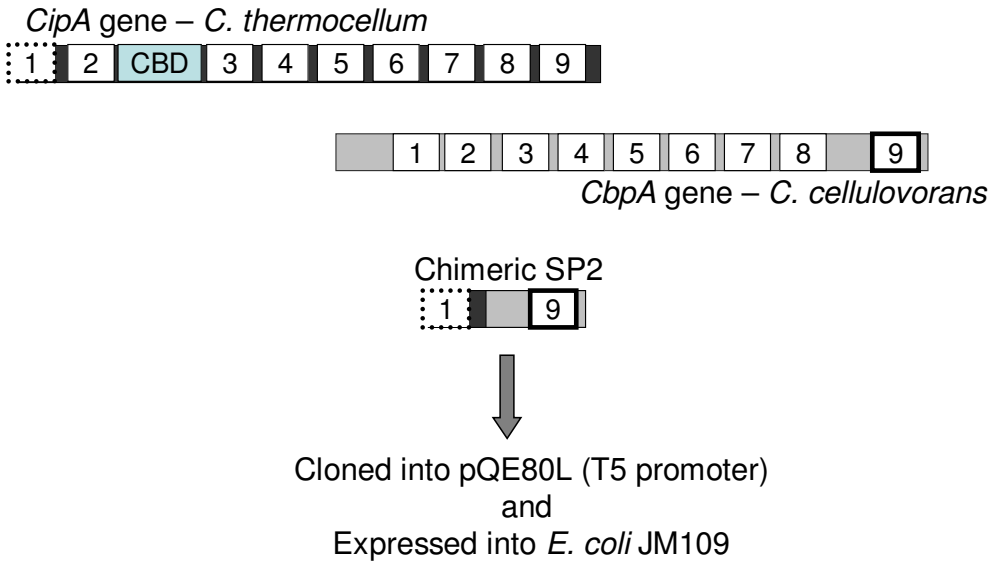


Figure 2.1 Schematic drawing of the construction of SP2 scaffolding protein.

The four enzymes are two xylanases (X10Doc and X11Doc), a ferulic acid esterase (FAEDoc), and a bi-functional arabinofuranosidase/xylosidase (ABFDoc). X10Doc is the mature protein from *XynC* of *C. thermocellum*, which encodes a family 10 glycosyl hydrolase with a family-22 carbohydrate-binding module for xylan binding and a dockerin domain [19]. Using gene fusion techniques, X11Doc was constructed by adding the dockerin domain of *XynC* to the C-terminus of a family 11 xylanase gene (*BH0899*) from *Bacillus halodurans* [20]. Similarly, FAEDoc is the result of fusing the ferulic acid esterase domain of *XynZ* of *C. thermocellum* [21] to a dockerin domain from *XynA* of *Clostridium cellulovorans* [22]. Lastly, a bi-functional arabinofuranosidase/xylosidase isolated from the metagenome of a compost starter mixture [23] was also functionalized with the dockerin

domain from *XynA* of *C. cellulovorans* to create ABFDoc. As shown in Figure 2.2, linker sequences were included in all enzymes between the catalytic and dockerin domains to ensure sufficient flexibility for proper folding of each segment and promote xylanosome construction. The physical characteristics for each xylanosomal enzyme are listed in Table 2.2. The molecular weights for each enzyme are within the range of enzymes currently found in natural and designer cellulosomes [24].

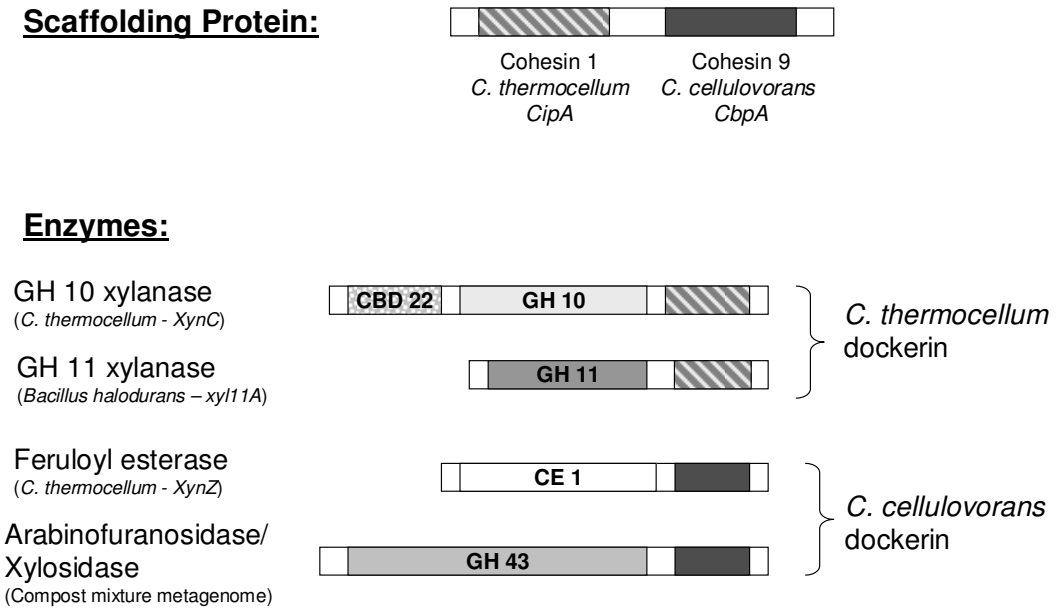


Figure 2.2 Modular schematic of gene construction for each xylanosome component.

Table 2.1 Physical characteristics of xylanosomal enzymes

Enzyme	**Molecular Weight (Da)**	**Optima**	**Origin of Dockerin Domain**	**Recombinant Strain**
X10Doc	70,575	pH 5.5, 80^{o}C	*C. thermocellum*	*E. coli* BL21Star (DE3)/ pEXynC
X11Doc	30,179	pH 7 , 40^{o}C	*C. thermocellum*	*E. coli* BL21Star (DE3)/ pQX11Doc
FAEDoc	38,652	pH 7 , 40^{o}C	*C. cellulovorans*	*E. coli* BL21Star (DE3)/ pEFAEDoc
ABFDoc	68,188	pH 6, 55^{o}C	*C. cellulovorans*	*E. coli* BL21Star (DE3)/ pEdeAXDoc

The optimized expression conditions for each protein varied between 0.2 – 1.0 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for induction and incubation at 16 – 30^{o}C for 5 – 24 hours induction time. Each protein was successfully purified using immobilized nickel(II) affinity chromatography via N-terminal (X10Doc, X11Doc, SP2) or C-terminal (FAEDoc, ABFDoc) histidine tags. Summary of purification results are outlined in Table 2.2. Since SP2 does not have enzymatic activity, specific activity cannot be calculated.

Table 2.2 Purification table for components of two-unit xylanosomes[a,b]

Xylanosome component	Activity (U)	Protein (mg)	Specific Activity (U/mg)	Purification Factor	% Recovery
X10Doc	65	0.5	127.3	383	48
X11Doc	26.6	0.35	75.5	13.1	50
FAEDoc	0.16	0.03	5.48	39.5	29
ABFDoc	0.66	0.18	3.7	7	6
SP2	n/a	2.2	n/a	n/a	12.4

[a] These values represent a typical purification using IMAC, from an average of two purifications.
[b] For ABFDoc, values listed are for arabinofuranosidase activity only.

2.3.2 Construction of two-unit xylanosomes

As shown in Figure 2.3, the two-unit scaffoldin (SP2) has two different cohesins from different species, allowing four different xylanosomes to be constructed from the four dockerin-containing hemicellulases. The xylanosome combinations are: X10Doc + FAEDoc, X10Doc + ABFDoc, X11Doc + FAEDoc, and X11Doc + ABFDoc.

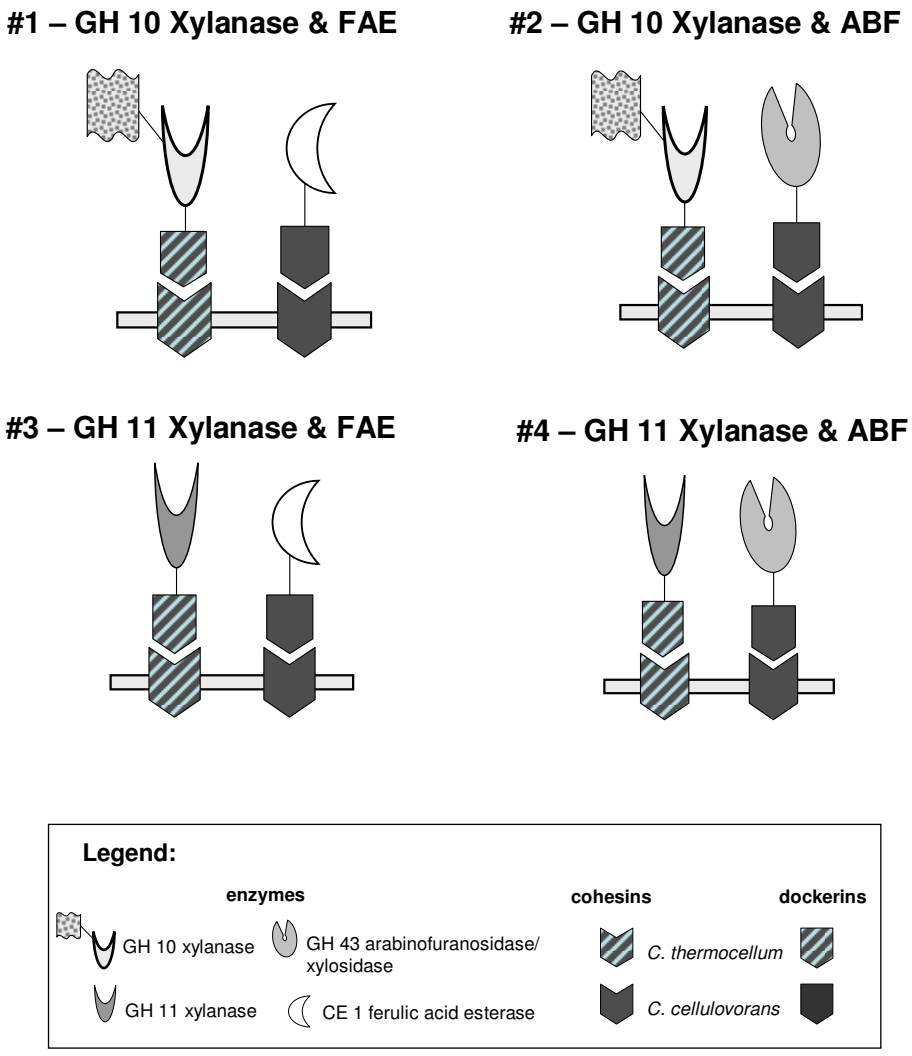


Figure 2.3 Composition of the four two-unit xylanosomes. ABF – ABFDoc; FAE – FAEDoc; GH 10 Xylanase – X10Doc; GH 11 Xylanase – X11Doc.

Based on its enzymatic makeup, each xylanosome has a different target in the degradation of xylan. All xylanosomes contain a xylanase, X10Doc or X11Doc, for degrading the xylan backbone of β-1,4-linked xylose molecules, as this step is crucial in the overall hydrolysis of the substrate. Those xylanosomes containing the bi-functional arabinofuranosidase/xylosidase, ABFDoc, are targeted for increased release of reducing sugars from xylan. FAEDoc-containing xylanosomes are designed for the release of ferulic acid, in addition to release of reducing sugars.

Surface plasmon resonance was used to confirm construction of the two-unit xylanosomes. As depicted in Figure 2.4, SPR consists of measuring changes in the refractive index (RI) of a prism lined with a gold surface. The RI of the prism varies according to changes in mass on top of the gold surface. Functionalization of the gold surface allows for immobilization of a protein. This added mass results in a shift to the overall RI of the prism, which is detectible by a spectrophotometer. Exposing an analyte that can interact with the immobilized protein continues to alter the RI of the prism by adding more mass to the prism surface. In this case, the SP2 scaffolding protein is immobilized on the gold surface and dockerin-tagged enzymes are the analytes. SPR will measure the protein-protein interaction between the cohesins of SP2 with the dockerins of the xylanosome enzymes. As dockerins bind with cohesins, the mass on the prism surface increases, thus causing changes in the RI of the prism. These changes are measured in micro-refractive index units (μRIU) over time.

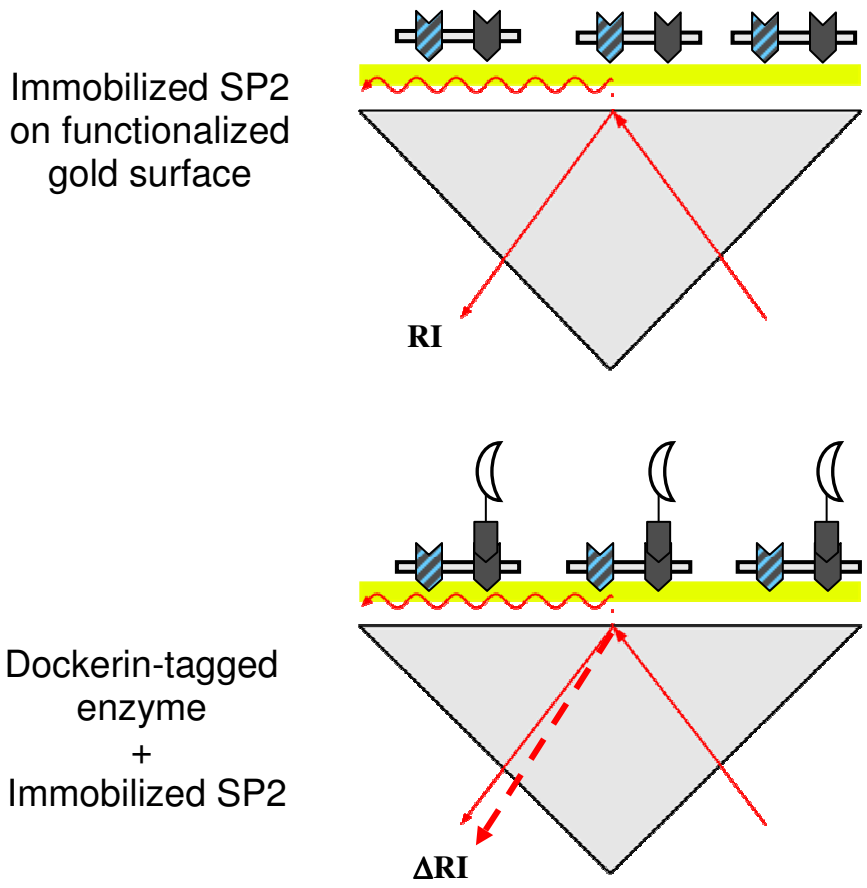


Figure 2.4 Schematic of mechanism of surface plasmon resonance.

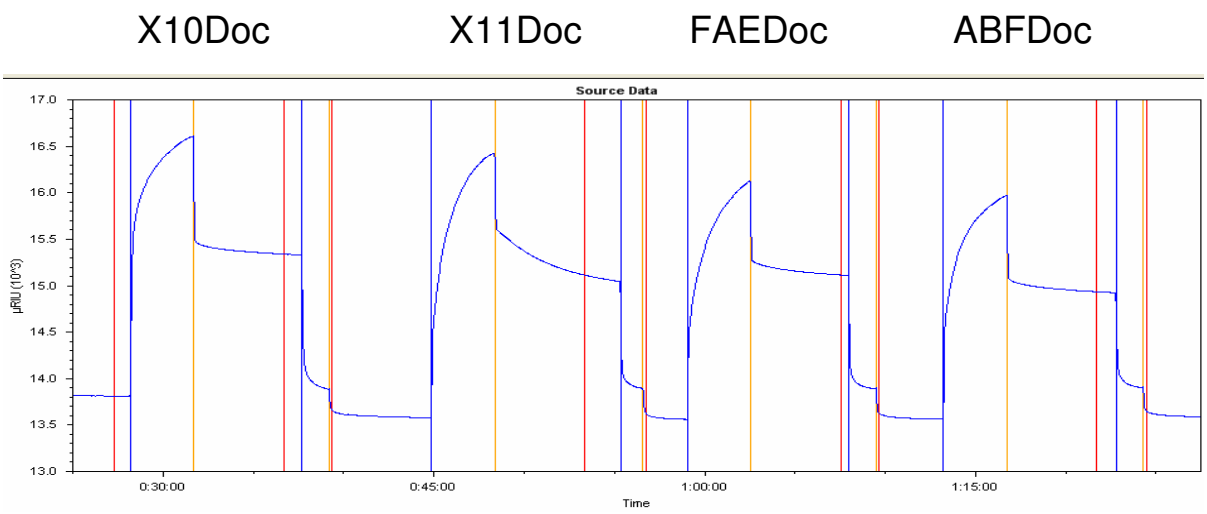


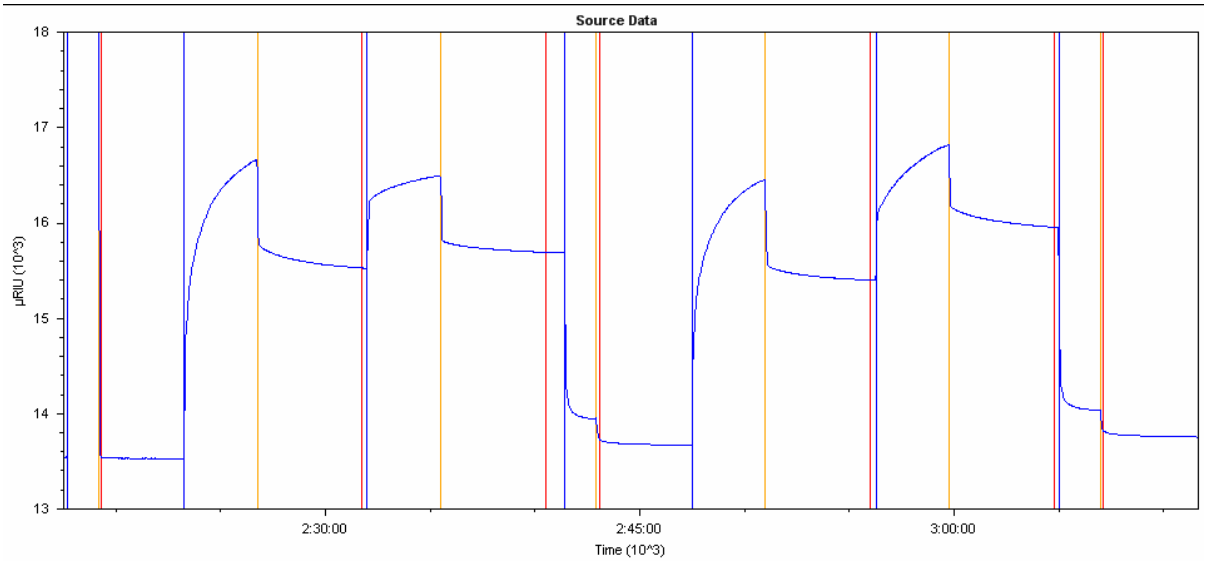
Figure 2.5 SPR spectra (micro-Refractive Index Units vs Time (h:mm:ss)) of individual dockerin-tagged enzymes interacting with immobilized SP2.

Once SP2 was immobilized, each dockerin-tagged enzyme was exposed to the surface. Spectra in Figure 2.5 verify that all enzymes interact with the immobilized SP2, signifying that each cohesin/dockerin system is functional. The µRIU signal at the beginning of each analyte injection gives a large positive response, representing an increase of mass on the prism surface; increased mass confirms proper protein-protein interaction between the cohesins of SP2 and dockerins of each enzyme. Initial exposure or injections of known concentrations of analyte over the immobilized protein provides the rate of association, k_{on}, for the analyte-protein complex. In Figure 2.5, this is represented as the µRIU signal increases from the baseline, or the signal between the blue and yellow vertical lines. Furthermore, flowing running buffer in the absence of analyte determines the rate of dissociation, k_{off}, of the analyte from the immobilized protein. This phenomenon is represented in Figure 2.5 as the signal that decreases immediately after the drastic signal increase, or the signal between the yellow and red vertical lines. From these values, an

affinity constant, K_A, defined as k_{on}/k_{off}, or dissociation constant, $K_D = k_{off}/k_{on}$, can be measured to gain insight into the strength of the protein-protein interaction between cohesins and dockerins. As a negative control, the signal representing the non-specific binding of the analyte to the gold surface is detected and subtracted from the final signal before analysis. The kinetics of the cohesin-dockerin interaction are discussed in Chapter 3.

Individual enzymes were loaded sequentially to ensure that the xylanosomes could form regardless of which enzyme bound initially. Figure 2.6 shows that all four xylanosomes form. After exposure of the first enzyme, either X10Doc or X11Doc, the overall μRIU signal is higher than with the immobilized SP2 protein only (baseline), indicating an increase in mass on the prism surface. This added mass represents a confirmed interaction between the cohesins and dockerins of *C. thermocellum*. When the second enzyme, either FAEDoc or ABFDoc, is subsequently exposed to the immobilized SP2, the overall μRIU signal again increases, signifying that the additional mass on the prism represents the proper function of the *C. cellulovorans* cohesin-dockerin system. From these spectra, we can confirm that the two-unit xylanosomes were successfully constructed. The prism surface was regenerated after each xylanosome formation to remove dockerin-tagged enzymes.

(a) X10Doc → ABFDoc X10Doc → FAEDoc



(b) X11Doc → ABFDoc X11Doc → FAEDoc

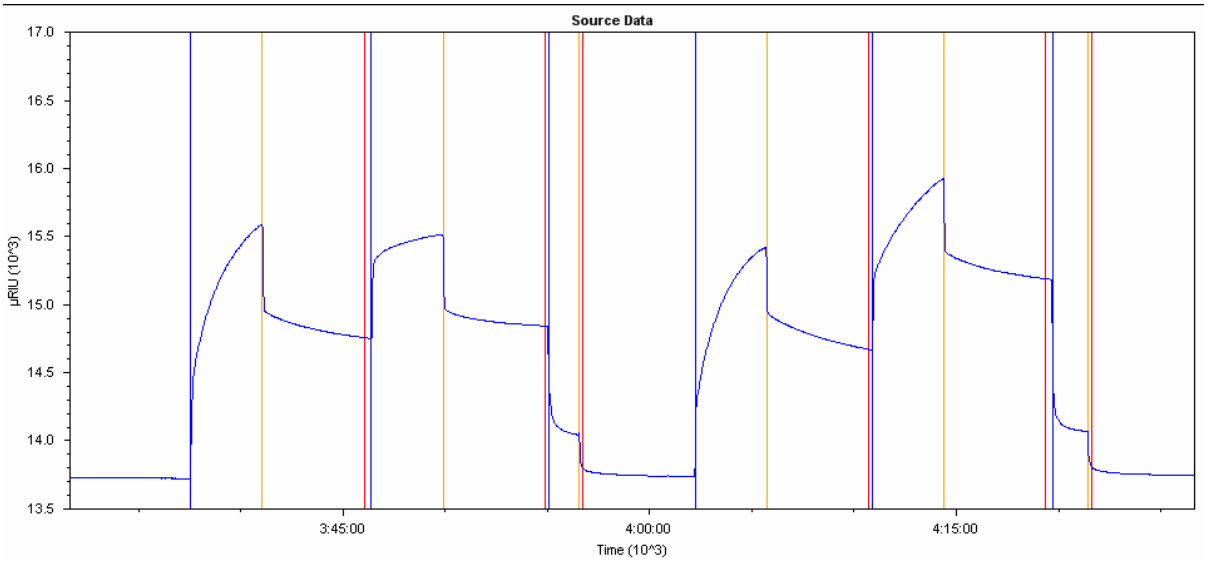
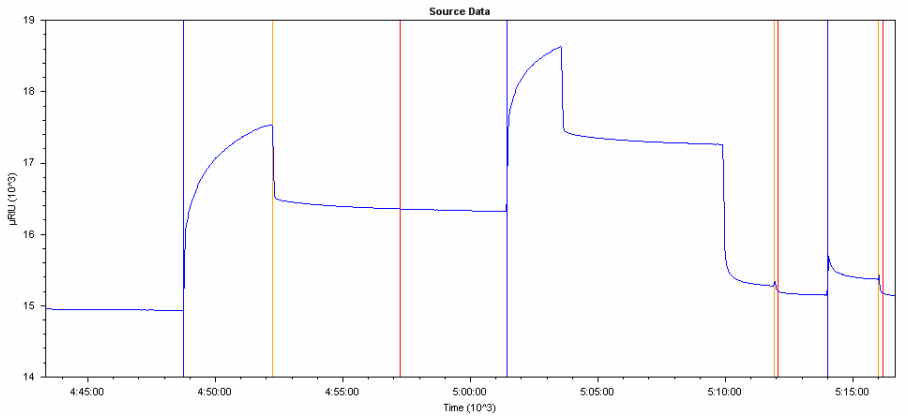


Figure 2.6 SPR spectra displaying the formation of two-unit xylanosomes. (a) Xylanosomes containing X10Doc and (b) xylanosomes containing X11Doc.

Conveniently, comparison of the formation of xylanosomes containing FAEDoc between Figures 2.6 and 2.7 shows that both xylanosomes form, whether a xylanase (*C. thermocellum* dockerin system) or the FAEDoc is added first or last because of the significant increase in μRIU after exposure of SP2 to each enzyme.

(a) FAEDoc → X10Doc



(b) FAEDoc → X11Doc

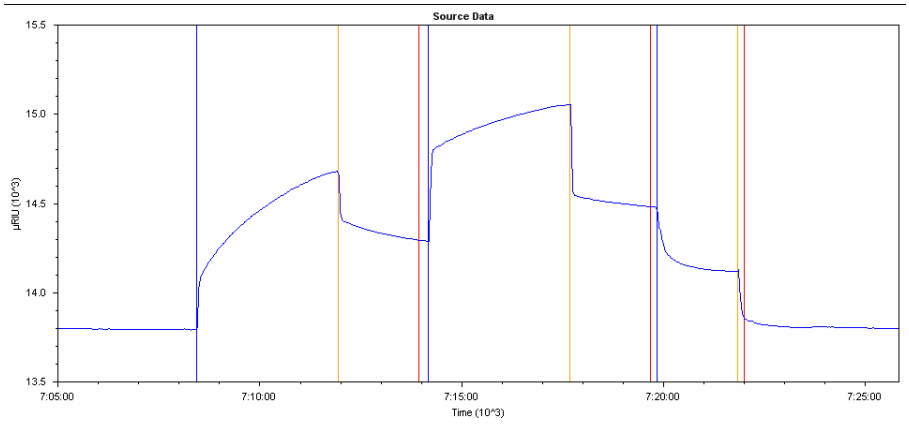
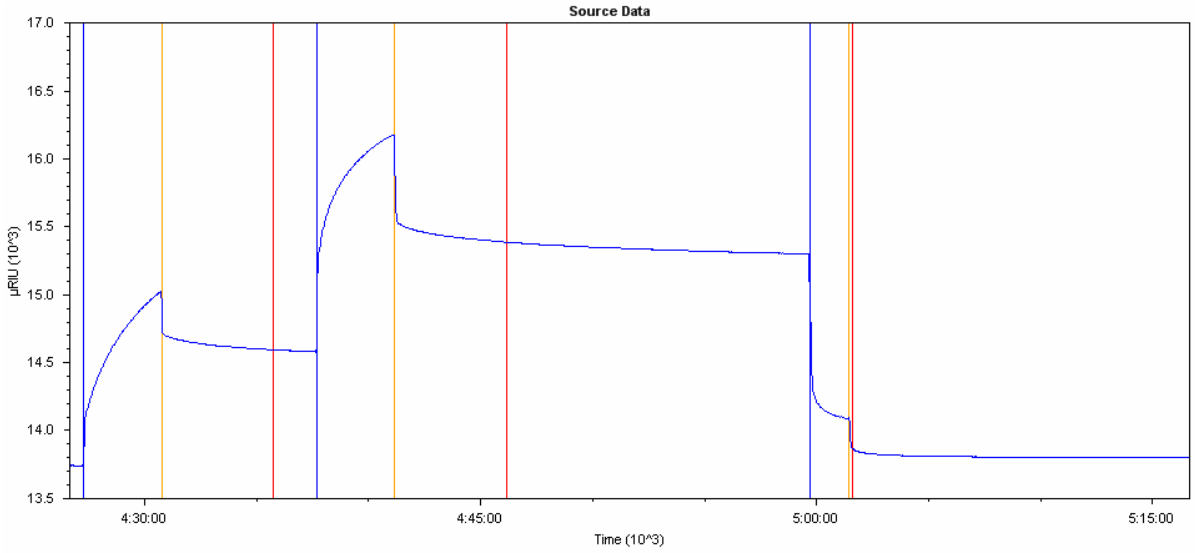


Figure 2.7 SPR spectra of the construction of xylanosomes with initial addition of FAEDoc. (a) Sequential addition with FAEDoc followed by X10Doc. (b) Sequential addition with FAEDoc followed by X11Doc.

However, for xylanosomes containing ABFDoc, spectra in Figure 2.6 do not show as large of a change in RI, indicating that ABFDoc does not interact with SP2 as well as FAEDoc. Figure 2.8 displays how changing the order of enzyme addition from X10Doc → ABFDoc to ABFDoc → X10Doc allows for proper formation of the two-unit xylanosome. A higher μRIU signal indicates more ABFDoc interacts with SP2 when it is exposed to SP2 first; the two-unit structure is formed when X10Doc is then added to the SP2—ABFDoc structures. However, changing the sequential order of the X11Doc + ABFDoc xylanosome does not drastically improve the formation of xylanosomes. The overall μRIU signal after exposure to both X11Doc and ABFDoc is similar with no matter the order of enzyme addition. Interestingly, both of these enzymes are non-cellulosomal, meaning that they are from proteins not found within cellulosomes. These results show ABFDoc should be added first, since fewer xylanosomes form when either X10Doc or X11Doc is present in the scaffoldin before ABFDoc. This information is important in preparing xylanosome solutions for biomass hydrolysis; it must be certain that the structures are constructed for proper comparison of hydrolysis to free enzyme systems.

(a) ABFDoc → X10Doc



(b) ABFDoc → X11Doc

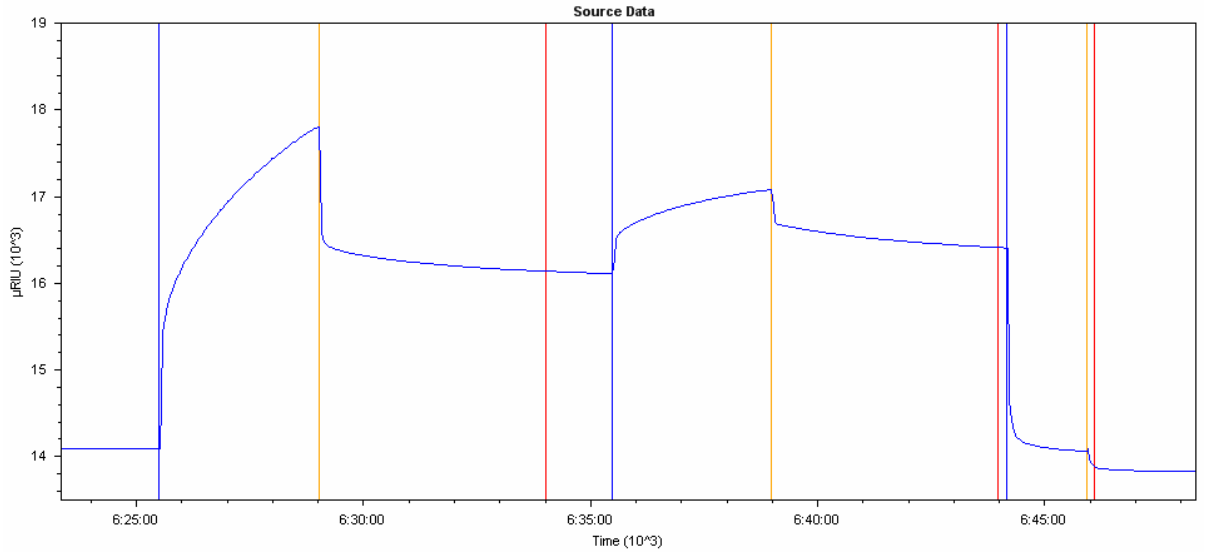


Figure 2.8 SPR spectra of the construction of xylanosomes with initial addition of ABFDoc. (a) Sequential addition with ABFDoc followed by X10Doc. (b) Sequential addition with ABFDoc followed by X11Doc.

2.4 Discussion

Enzymatic components of the two-unit xylanosomes were selected based on their ease of expression using *E. coli* systems and physical characteristics, such as pH and temperature optima, that allow for synergistic action in the degradation of xylan. More importantly, the selected genes encoded enzymes that were monomeric in their quaternary structure, ensuring a 1:1 ratio of dockerin to cohesin interaction and proper construction of the xylanosome biocatalyst. All selected components were able to be cloned and expressed into *E. coli* and purified using IMAC via N- or C-terminal histidine tags. Surface plasmon resonance confirmed interaction of individual enzymes with the scaffolding protein, as well as formation of the all four two-unit xylanosomes. Interestingly, xylanosomes containing FAEDoc could form no matter the sequential order of the enzymes to the scaffolding protein. However, the remaining two structures required that ABFDoc was to be added first for proper formation of the xylanosomes. The X11Doc + ABFDoc structure was the most difficult to form using SPR. Both X11Doc and ABFDoc are non-cellulosomal; therefore, folding or orientation of the catalytic domains compared to the dockerin domains may not be optimal for these particular enzymes. Overall, these biocatalysts are now available for application in xylan hydrolysis, specifically in the release of reducing sugars and/or ferulic acid into solution from the insoluble xylan substrate.

2.5 Materials and Methods

2.5.1 Strains and materials

Clostridium cellulovorans 743B ATCC 35296 and *Bacillus halodurans* ATCC 21591 strains were purchased from American Type Cell Culture (ATCC, Virginia, U.S.A.).

Bacterial Genomic DNA Extraction Kits from Sigma-Aldrich (U.S.A) were used for isolating the genomic DNA of *C. cellulovorans* and *B. halodurans*. The genomic DNA from *Clostridium thermocellum* ATCC 27405 was purchased from ATCC. *E. coli* TOP10 and BL21Star (DE3) strains were obtained from Invitrogen (U.S.A.); *E. coli* JM109 was already present in our lab strain collection. Expression vectors pET20b(+) and pET160-TOPO® containing T7 promoters were obtained from Novagen (California, U.S.A.) and Invitrogen (California, U.S.A.), respectively. The expression plasmid pQE80L was collected from Qiagen (California, U.S.A.). The *E. coli* strain harboring the plasmid containing the bi-functional arabinofuranosidase/xylosidase, pET22b(+)-deAX, from the metagenome of a compost starter mixture [23] was graciously provided by Kurt Wagschal from the United States Department of Agriculture. All restriction enzymes were obtained from New England Biolabs.

2.5.2 Cloning and expression of X10Doc

The mature protein sequence of *XynC* was amplified via polymerase chain reaction (PCR) from the genomic DNA of *Clostridium thermocellum* ATCC 27405 using iProof, a high-fidelity proofreading DNA polymerase (Bio-Rad) [19]. The primers used, XynC-F and XynC-R, are listed in Table 2.1. The forward primer contains a 5'-CACC overhang for the proper insertion into the pET160-TOPO® vector for directional TOPO cloning. A single band PCR product was confirmed by agarose gel electrophoresis by visualizing a single band at approximately 1800 bp. The purified DNA product was then ligated into the pET160-TOPO® vector and transformed into One Shot *E. coli* TOP10 competent cells according to the vector manual. A minimum of five recombinant colonies were screened for the presence of the pEXynC (7.6 kb) plasmid by first isolating the plasmid using a plasmid extraction kit

from an overnight culture in Luria Broth (LB) containing 100 μg/mL ampicillin (LB/amp100). Plasmids were purified using either ZymoResearch or Qiagen mini-prep kits according to the provided manuals and digested using *Hpa*I to confirm for proper insertion of *XynC*. The insert sequence and open reading frame of *XynC* was verified with DNA sequencing. pEXynC was maintained in *E. coli* TOP10 as expression plasmids were not sustained well in the expression strain, *E. coli* BL21Star (DE3). After chemical transformation, *E. coli* BL21Star (DE3) harboring pEXynC was cultured overnight in 10 mL LB/amp100 at 37°C, 250 rpm. The seed culture (5 mL) was used to inoculate 100 mL LB/amp100, grown to an optical density of 0.4 at 37°C, 250 rpm. The culture was then induced with 100 mM isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM and incubated at 25°C for 16 hours. Cells were harvested by centrifugation at 3500 x *g* for 30 minutes at 4°C and stored at -20°C until needed. For confirmation of protein expression, cells were re-suspended to 10X concentrations in either 10 mM Tris-HCl, pH 7.5 or 50 mM sodium phosphate, 0.3 M NaCl, pH 8.0, followed by sonication at 20s bursts with 1 minute rest on ice and repeated six to seven times. Crude lysates were collected by centrifugation at 4000 x *g* for 40 minutes at 4°C and removing the supernatant. Expression of the target protein in the lysates was confirmed using SDS-PAGE with observation of the appearance of a protein band at 70 kDa, using the soluble fraction of a non-induced sample as a control, and stained with Bio-safe Coomassie Blue (Bio-Rad) for band visualization. Enzymatic activity was confirmed on oat spelt xylan as well. All protein concentrations were measured via Bradford assay using a protein reagent dye (Bio-Rad), with bovine serine albumin as the standard.

Table 2.3 List of primers for cloning of xylanosome components

Component/Gene Source	**Primer Set (5' → 3')**
Dockerin domain/ *XynA*, *Clostridium cellulovorans*	CcDoc-F: CTTAGTCGACACCCCACCTGTTATCGTAGACCAGCCTACC CCAGAA CcDoc-R: GTGTCTCGAGGTTTGGTTCTTCAGATGTAGTTGG
Dockerin domain/ *XynC*, *Clostridium thermocellum*	CtDoc-F: GACCTAAGCTTGTTAACAGTGTTCCGCCGC CtDoc-R: GCACAGTCGACAAGTTCTCTCAGGACGAGTTTTTTC
Xylanase, GH 10/ *XynC, Clostridium thermocellum*	XynC-F: CACCGCAGCTCTGATTTACGATGATTT XynC-R: AGACATGCAAGCACAACGG
Xylanase, GH 11/ *BH0899*, *Bacillius halodurans*	Xyn11A-F: GAGTCAGGATCCATGAATACCTACTGGCAATACTGGACC Xyn11A-R: TACACTAAGCTTCCAAACCGTCACATTCGATC
Ferulic acid esterase, CE 1/ *XynZ*, *Clostridium thermocellum*	FAE-F: CATCAGGATCCGCTAGTCACAATAAGCAGTACATCAG FAE-R: TCTTCGTCGACAGTGTTTCCATCCCTCGTCAA
Cohesin 1/ *CipA*, *Clostridium thermocellum*	CtDoc-F: CCCTTTAGATCTGGTGGCGCAGCTATGATACCGCAGACA GTAT CtDoc-R: GCAGTTTTTTCAAATGTTGCTACTACTGGAGTTATTGTAC CACCGTCTGACGGAACATTTG
Cohesin 9/ *CbpA*, *Clostridium cellulovorans*	CcDoc-F: TACCGACAACACAGCCAAATGTTCCGTCAGACGGTGGTA CAATAACTCCAGTAGTAGCAAC CcDoc-R: TTTGAGCTCTTAGCTAACTTTAACACTTC

2.5.3 Cloning and expression of X11Doc

The mature *Xyn11A* gene was isolated from the genomic DNA of *Bacillus halodurans* via PCR using the primers listed in Table 2.1 [20]. Forward and reverse primers, Xyn11A-F and Xyn11A-R, contain *Bam*HI and *Hind*III restriction sites, respectively. The genomic DNA of *Clostridium thermocellum* was used to amplify the dockerin domain of the *XynC* gene with the primers, CtDoc-F (*Hind*III) and CtDoc-R (*Sal*I). pQE80L was linearized with *Bam*HI and *Sal*I double digestion and ligated with *Bam*HI-*Hind*III-digested Xyn11A and *Hind*III-*Sal*I-digested CtDoc DNA fragments using T4 DNA ligase (New England Biolabs). Chemically-competent cells of *E. coli* JM109 were transformed using the described ligation mixture. Colony PCR was performed on recombinant colonies to confirm the presence of the Xyn11A-CtDoc fragment using Xyn11A-F and CtDoc-R primers. Plasmids were isolated from positive clones and screened for presence and proper size of pQX11Doc (5.5 kb) using agarose gel electrophoresis. Expression of X11Doc was carried out by culturing positive clones in 10 mL LB/amp100 media at 37^{o}C, 250 rpm for overnight and inoculating 100 mL LB/amp100 with 5 mL of the seed culture. Once the cells reached an OD between 0.4 – 0.5, 100 mM IPTG was added to a final concentration of 0.5 mM and the culture incubated at 25^{o}C, 250 rpm for 16 hours. Cells were harvested by centrifugation at 3500 x *g* for 30 minutes at 4^{o}C and stored at -20^{o}C until needed. Confirmation of protein expression was carried out in a manner similar to that described in Section 2.3.2, with an expected protein band at 30 kDa in crude lysates compared to band pattern from non-induced sample as a control, as well as activity on oat spelt xylan.

2.5.4 Cloning and expression of FAEDoc

The ferulic acid esterase (FAE) domain from *XynZ* was cloned from *C. thermocellum* genomic DNA using FAE-F (*Bam*HI) and FAE-R (*Sal*I) primers [21]. Genomic DNA from *C. cellulovorans* was used to amplify the dockerin domain from the *XynA* gene with CcDoc-F and CcDoc-R primers [22]. Both primer sets are listed in Table 2.1. Single product bands for both FAE and CcDoc fragments were confirmed using agarose gel electrophoresis. FAE was double-digested with *Bam*HI and *Sal*I; CcDoc was digested with *Sal*I and *Xho*I. The digested fragments were inserted into *Bam*HI-*Xho*I linearized pET20b(+) to generate pEFAEDoc. The subsequent ligation mixture used to transform *E. coli* TOP10 for vector confirmation and propagation. Recombinant clones were screened for proper plasmid size (4.7 kb), as well as proper orientation of the FAEDoc chimeric gene via PCR using the primers FAE-F and CcDoc-R. Expression of FAEDoc was carried out as mentioned in Section 2.3.2, with induction conditions of 1 mM IPTG at 30^{o}C for 5 hours. Protein expression was carried out in a manner similar to that described in Section 2.3.2, with an expected protein band at approximately 39 kDa compared to band pattern from non-induced sample as a control and confirmed via activity methyl ferulate.

2.5.5 Cloning and expression of ABFDoc

The *C. cellulovorans* dockerin, CcDoc, as described in the previous section was digested with *Sal*I and *Xho*I for ligation into *Xho*I-linearized pEdeAX. *E. coli* TOP10 was transformed with the ligation mixture and recombinant colonies screened for presence of the pEdeAXDoc plasmid (7.5 kb) via plasmid size. Correct orientation of the dockerin domain was confirmed by digesting isolated plasmids with *Sal*I. Expression of

ABFDoc was performed as described in Section 2.3.2, except the induction conditions were 0.2 mM final IPTG concentration and incubation at 16°C for 24 hours. Protein expression was carried out in a manner similar to that described in Section 2.3.2, with an expected protein band at approximately 68 kDa compared to band pattern from non-induced sample as a control and confirmed via presence of activity on *p*-nitrophenyl α-L-arabinofuranoside and *p*-nitrophenyl β-D-xylopyranoside.

2.5.6 Cloning and expression of SP2

Using the primers SP2-F and SP2-R listed in Table 2.1, the first type I cohesin within the *CipA* gene of *C. thermocellum* and the ninth type I cohesin from the *CbpA* gene of *C. cellulovorans* was isolated via PCR with the respective genomic DNA from each organism. Each PCR product was purified and subsequently used to perform an overlap PCR to fuse the chimeric scaffolding protein sequence. The resulting SP2 fragment was digested with *Bgl*II and *Hind*III for ligation into *Bam*HI-*Hind*III linearized pQE80L. Recombinant pQH1H9 was used to transform *E. coli* JM109 for vector propagation and protein expression. Presence of the target plasmid was confirmed as mentioned in previous sections. Expression of SP2 was performed as outlined in Section 2.3.3, except 1 mM IPTG was used for induction at 30°C for 5 hours.

2.5.7 Purification of xylanosome components

Crude lysates for each expressed protein were prepared as mentioned in Section 2.3.2, using 50 mM sodium phosphate, 0.3M NaCl buffer, pH 8.0. Each target protein was purified using immobilized metal (nickel (II)) affinity chromatography (HIS-Select™ Resin, Sigma-Aldrich, U.S.A.). Lysates were loaded onto resin equilibrated

with 50 mM sodium phosphate, 0.3M NaCl buffer, pH 8.0, in batch mode on ice overnight to bind the target proteins. Next, the resin was washed extensively in either batch or column mode between 30 – 50 times the resin volume with the same buffer to remove non-specifically bound proteins. Bound target proteins were eluted from the resin with 50 mM sodium phosphate, 0.3 M NaCl, 250 mM imidazole buffer, pH 8.0, and the elute fraction dialyzed against 10 mM Tris-HCl, pH 7.5 for removal of imidazole for subsequent assays. To confirm purification, the crude, non-binding, and elute fractions were analyzed via SDS-PAGE using either a 12% Tris gel or Any kD™ gradient gel from Bio-Rad (U.S.A.), running at 120 V for 45 – 60 minutes or 200 V for 30 minutes, respectively and stained with Bio-safe Coomassie Blue (Bio-Rad) for band visualization.

2.5.8 Enzymatic activity assays

Xylanase activity of X10Doc and X11Doc was measured by adding 50 µL of appropriately diluted enzyme to 50 µL of 1% (w/v) oat spelt xylan (Sigma-Aldrich, U.S.A.) in 100 mM sodium phosphate, pH 7.0 and incubating in a 60^{o}C water bath for 10 minutes. Released reducing sugars from oat spelt xylan were measured using the dinitrosalicyclic (DNS) acid method, using xylose as a standard [25]. 1 unit (U) of activity is denoted by the amount of enzyme to release 1 µmol of reducing sugar per minute.

Arabinofuranosidase and xylosidase activities for ABFDoc were performed by mixing 100 µL of appropriately diluted enzyme solution with either 1 mM *p*-nitrophenyl alpha-L-arabinofuranoside or *p*-nitrophenyl beta-D-xylopyranoside (Sigma-Aldrich, U.S.A.) in 100 mM phosphate buffer, pH 6.0, to a final volume of 1 mL. The reaction mixture was incubated in a water bath at 50^{o}C for 10 minutes and the absorbance at 410

nm immediately measured. Activity units are measured by the amount of enzyme required to release 1 μmol of *p*-nitrophenol (Sigma-Aldrich, U.S.A.) per minute. Calibration curves were generated using *p*-nitrophenol as the standard.

Ferulic acid esterase activity of FAEDoc was determined by measuring the amount of ferulic acid released from methyl ferulate (Apin Chemicals, Oxon, UK). Reactions were carried out by adding 50 μL of enzyme solution to 440 μL of 100 mM sodium phosphate buffer, pH 7.0, with 10 μL of 50 mM methyl ferulate (1 mM final concentration) and incubating the mixture at 40^{o}C for 30 minutes. The reaction was stopped by adding 200 μL of glacial acetic acid. Quantification of released ferulic acid was performed by high performance liquid chromatography on an Agilent 1100 system with a Li-Chrosopher RP-18 column with a gradient of 0.01% (v/v) acetic acid and methanol for the mobile phase [26]. One unit of activity is defined as the amount of enzyme that releases 1 μmol of ferulic acid per min of reaction. Calibration curve was generated using ferulic acid (Apin Chemicals, UK).

2.5.9 Surface plasmon resonance

Experiments were performed using a Reichert system (Reichert, New York, U.S.A.), with a running buffer of either Phosphate-Buffered Saline with Tween (PBST, 10 mM sodium phosphate, 150 mM NaCl, 0.05% (v/v) Tween 20, pH 7.2) for scaffolding immobilization or 20 mM Tris-maleate, 1 mM CaCl2, 0.05% (v/v) Tween 20, pH 6.0. SP2 was immobilized onto a SR700 gold sensor slide with a mixed self-assembled monolayer of 90% OH-$(PEG)_6$-C_{11}-SH/10% COOH-$(PEG)_6$-C_{11}-SH. The slide was functionalized using 0.2M N-(3-Dimethylaminopropyl)-N'-ethylcarbodiimide (EDC, Sigma-Aldrich, U.S.A.) and N-hydroxysuccinimide (NHS, Sigma-Aldrich, U.S.A.), then

exposed to SP2 (20 μg/mL) in 10 mM sodium acetate, pH 4.4, followed by 1M ethanolamine to cap any unreacted sites. Dockerin-tagged enzymes were diluted with Tris-maleate running buffer to 100 nM concentration and allowed to interact with the immobilized SP2 for 180-s injections. The slide surface was regenerated using 15 mM HCl for 90 to 120 seconds.

2.6 References

1. Jindou, S., et al., *Cohesin-dockerin interactions within and between Clostridium josui and Clostridium thermocellum: binding selectivity between cognate dockerin and cohesin domains and species specificity.* J Biol Chem, 2004. **279**(11): p. 9867-74.

2. Mechaly, A., et al., *Cohesin-dockerin interaction in cellulosome assembly: a single hydroxyl group of a dockerin domain distinguishes between nonrecognition and high affinity recognition.* J Biol Chem, 2001. **276**(13): p. 9883-8.

3. Fierobe, H.P., et al., *Degradation of cellulose substrates by cellulosome chimeras. Substrate targeting versus proximity of enzyme components.* J Biol Chem, 2002. **277**(51): p. 49621-30.

4. Fierobe, H.P., et al., *Action of designer cellulosomes on homogeneous versus complex substrates: controlled incorporation of three distinct enzymes into a defined trifunctional scaffoldin.* J Biol Chem, 2005. **280**(16): p. 16325-34.

5. Mingardon, F., et al., *Incorporation of fungal cellulases in bacterial minicellulosomes yields viable, synergistically acting cellulolytic complexes.* Appl Environ Microbiol, 2007. **73**(12): p. 3822-32.

6. Mingardon, F., et al., *Exploration of new geometries in cellulosome-like chimeras.* Appl Environ Microbiol, 2007. **73**(22): p. 7138-49.

7. Mingardon, F., et al., *Heterologous production, assembly, and secretion of a minicellulosome by Clostridium acetobutylicum ATCC 824.* Appl Environ Microbiol, 2005. **71**(3): p. 1215-22.

8. Vazana, Y., et al., *Interplay between Clostridium thermocellum family 48 and family 9 cellulases in cellulosomal versus noncellulosomal states.* Appl Environ Microbiol. **76**(10): p. 3236-43.

9. Arai, T., et al., *Synthesis of Clostridium cellulovorans minicellulosomes by intercellular complementation.* Proc Natl Acad Sci U S A, 2007. **104**(5): p. 1456-60.

10. Fierobe, H.P., et al., *Cellulosome from Clostridium cellulolyticum: molecular study of the Dockerin/Cohesin interaction.* Biochemistry, 1999. **38**(39): p. 12822-32.

11. Haimovitz, R., et al., *Cohesin-dockerin microarray: Diverse specificities between two complementary families of interacting protein modules.* Proteomics, 2008. **8**(5): p. 968-79.

12. Cirino, P.C., J.W. Chin, and L.O. Ingram, *Engineering Escherichia coli for xylitol production from glucose-xylose mixtures.* Biotechnol Bioeng, 2006. **95**(6): p. 1167-76.

13. Dien, B.S., N.N. Nichols, and R.J. Bothast, *Recombinant Escherichia coli engineered for production of L-lactic acid from hexose and pentose sugars.* J Ind Microbiol Biotechnol, 2001. **27**(4): p. 259-64.

14. Ohta, K., et al., *Genetic improvement of Escherichia coli for ethanol production: chromosomal integration of Zymomonas mobilis genes encoding pyruvate decarboxylase and alcohol dehydrogenase II.* Appl Environ Microbiol, 1991. **57**(4): p. 893-900.

15. Yanase, H., et al., *Genetic engineering of Zymobacter palmae for production of ethanol from xylose.* Appl Environ Microbiol, 2007. **73**(8): p. 2592-9.

16. Deng, W., et al., *Variation of xylanosomal subunit composition of <i>Streptomyces olivaceoviridis</i> by nitrogen sources.* Biotechnology Letters, 2005. **27**(6): p. 429-433.

17. Jiang, Z.Q., et al., *Characterization of a novel, ultra-large xylanolytic complex (xylanosome) from Streptomyces olivaceoviridis E-86.* Enzyme and Microbial Technology, 2005. **36**(7): p. 923-929.

18. Jiang, Z.-Q., et al., *A novel, ultra-large xylanolytic complex (xylanosome) secreted by <i>Streptomyces olivaceoviridis</i&gt.* Biotechnology Letters, 2004. **26**(5): p. 431-436.

19. Hayashi, H., et al., *Sequence of xynC and properties of XynC, a major component of the Clostridium thermocellum cellulosome.* J Bacteriol, 1997. **179**(13): p. 4246-53.

20. Wamalwa, B.M., et al., *High-level heterologous expression of Bacillus halodurans putative xylanase xyn11a (BH0899) in Kluyveromyces lactis.* Biosci Biotechnol Biochem, 2007. **71**(3): p. 688-93.

21. Blum, D.L., et al., *Feruloyl esterase activity of the Clostridium thermocellum cellulosome can be attributed to previously unknown domains of XynY and XynZ.* J Bacteriol, 2000. **182**(5): p. 1346-51.

22. Kosugi, A., K. Murashima, and R.H. Doi, *Xylanase and acetyl xylan esterase activities of XynA, a key subunit of the Clostridium cellulovorans cellulosome for xylan degradation.* Appl Environ Microbiol, 2002. **68**(12): p. 6399-402.

23. Wagschal, K., et al., *Biochemical characterization of a novel dual-function arabinofuranosidase/xylosidase isolated from a compost starter mixture.* Appl Microbiol Biotechnol, 2009. **81**(5): p. 855-63.

24. Fontes, C.M. and H.J. Gilbert, *Cellulosomes: highly efficient nanomachines designed to deconstruct plant cell wall complex carbohydrates.* Annu Rev Biochem, 2010. **79**: p. 655-81.

25. Miller, G.L., *Use of dinitrosaliclic acid reagen for determination of reducing sugars.* Analytical Chemistry, 1959(31): p. 127-132.

26. Shin, H.D. and R.R.Z. Chen, *Production and characterization of a type B feruloyl esterase from Fusarium proliferatum NRRL 26517.* Enzyme and Microbial Technology, 2006. **38**(3-4): p. 478-485.

CHAPTER 3

CHARACTERIZATION AND APPLICATION OF TWO-UNIT XYLANOSOMES

3.1 Abstract

Hemicellulose is the second most abundant renewable resource in lignocellulosic biomass after cellulose. Commonly used acidic pretreatments can effectively remove hemicellulose from biomass but reduce the overall sugar polymers available for hydrolysis, ultimately leaving biomass underutilized in bio-based chemical production. Using the cellulosomes of efficient cellulolytic bacteria as a template, four different self-assembling protein nanostructures called xylanosomes were used to improve the hydrolysis of xylan, a common type of hemicellulose. Each xylanosome was composed of two hemicellulases and tested on wheat arabinoxylan or destarched corn bran for enzymatic hydrolysis. After a 24 hour-incubation, soluble sugars released from arabinoxylan increased up to 30% with xylanosomes containing a xylanase and bi-functional arabinofuranosidase/xylosidase over the corresponding free, unstructured enzymes. Furthermore, xylanosomes with a xylanase and ferulic acid esterase removed between 15 – 20% more ferulic acid from wheat arabinoxylan. Xylanosomes also showed synergy with cellulases on destarched corn bran, suggesting a possible role in increasing access to cellulose during enzymatic hydrolysis. Possible application of two-unit xylanosomes for consolidated bioprocessing was determined by characterizing the cohesin/dockerin assembly mechanism at various temperatures, as well as in the presence of non-specific proteins and ethanol. Cell surface display of the scaffolding protein on

Escherichia coli allowed membrane-bound formation of xylanosomes that retained enzymatic activity on synthetic substrates.

3.2 Introduction

Enzymatic hydrolysis of cellulose and hemicellulose within biomass into fermentable sugars is the most expensive and time-intensive aspect of biofuel production. Strong acid or alkaline processes are capable of complete saccharification pretreated biomass [1]. However, enzymes are a more favorable means of hydrolysis due to direct use of hydrolysates in fermentations and elimination of economically-unfavorable neutralizing or detoxification steps. Nature provides two main pathways for utilizing biomass, one each from aerobic and anaerobic microbes. Since oxygen and therefore energy is plentiful in aerobic conditions, fungi and bacteria secrete copious quantities of biomass hydrolyzing enzymes into the extracellular environment. Companies like Novozymes and Genencor take advantage of cellulolytic aerobes by producing optimized enzyme cocktails from secretomes of filamentous fungi grown on various biomass substrates for carbon sources [2, 3]. Conversely, anaerobic microbes experience limited amounts of energy and consequently maximize enzyme usage by strategically directing cellulases and hemicellulases into protein nanostructures called cellulosomes [4]. Biomass hydrolysis via cellulosomes for biofuel production could be beneficial by requiring fewer enzymes without losing effectiveness, therefore, resulting in cheaper protein production costs.

The most well-known cellulosome-producing microbe, *Clostridium thermocellum*, is highly-efficient cellulolytic bacteria and its cellulosomes are currently being studied for optimal enzyme ratios with various biomass substrates [4]. Self-

assembly and cell-anchoring abilities of cellulosomes are provided by cohesins present within scaffolding (type I) or membrane-associated (type II) proteins and dockerins included on numerous biomass-hydrolyzing enzymes (type I) and select scaffolding proteins (type II) [5, 6]. Cohesins and dockerins from multiple cellulosome-producing bacteria have been instrumental in the development of designer cellulosomes. Taking advantage of their species-specificity, structures with known enzymatic composition were constructed by selectively tagging cellulases with different dockerins from *Clostridium thermocellum*, *Clostridium cellulolyticum* and *Ruminococcus flavefaciens* and building a scaffolding protein with corresponding cohesin domains [7-10]. Research performed with designer cellulosomes has led to a better understanding of how cellulosomes work, such as allowing close-proximity of enzymes to substrates and increasing synergy between enzymes [11, 12]. In most cases, placing cellulases in structured systems (cellulosome) allowed greater release of reducing sugars compared to the unstructured, free enzyme systems (aerobic) [11-14].

Studies done with designer cellulosomes provide a platform for developing novel, self-assembling protein systems with different functions. While designer cellulosomes target insoluble cellulose substrates, design and application of a protein nanostructure targeted toward hemicellulose has yet to be investigated. Xylan is the most abundant type of hemicellulose available in targeted biomass, and thus, an efficient method for enzymatic hydrolysis of xylan could result in higher yields of fermentable sugars and other valuable acids from such biomass. With greater sugar recovery, higher quantities of biofuel or value-added chemicals can be produced at biorefineries.

The work in this chapter describes the construction of four different protein nanostructures, termed xylanosomes, and their application in the hydrolysis of wheat arabinoxylan and destarched corn bran, specifically the release of soluble sugars and ferulic acid. Furthermore, two-unit xylanosomes are displayed on the cell-surface on *Escherichia coli* for broader applications within biomass utilization. Thus, the interaction of cohesins and dockerins from *C. thermocellum* and C. *cellulovorans* are investigated at various conditions to prove applicability in consolidated bioprocessing processes.

3.3 Results

3.3.1 Characterization of xylanosome components

3.3.1.1 Kinetic characterization

Enzymatic assays revealed that all chimeric enzymes retained activity after the addition of the dockerin domain, as shown in Table 3.1. However, the dockerin domain negatively affected the enzyme efficiency. While the K_M values remained similar, the V_{max} values decreased when the dockerin domain was added to the catalytic domain for all dockerin-tagged enzymes. Similar trends were observed when dockerin domains were removed from cellulosomal hemicellulases from *Neocallimastix frontalis* [15]. The maximum reaction rates were converted from U/mg to U/mmol to account for the differences in protein molecular weight. The most significant change was observed with ferulic acid esterase (FAE) with a 3.2-fold decrease in reaction rate with the presence of a dockerin domain. Bi-functional arabinofuranosidase/beta-xylosidase (ABF), xylanase family 11, and xylanase family 10 also showed a 2.9-, 2.2-, and 1.4-fold decrease in reaction rate, respectively, when dockerins were attached to the catalytic domain.

Interestingly, both xylanases gave the lowest change in maximum activity, suggesting that the folding or orientation of the catalytic domains can better accommodate additional non-catalytic domains. Also, protein size appears to have no influence on the lowering of V_{max} due to the dockerin domain; the selected proteins range from 30 to 70 kDa with and 20 – 59 kDa without the dockerin domain. Lastly, the catalytic domains of X10Doc and FAEDoc are both found in the *C. thermocellum* cellulosome, but FAE was hindered more than X10, suggesting that native location of the enzyme does not predict how enzyme efficiency changes with the addition of dockerin domains. The degree of loss of enzyme activity appears to be more dependent on the enzyme action, i.e. cleavage of glycosyl linkages (xylanases) versus ester linkages (ferulic acid esterase).

Table 3.1 Kinetic parameters for enzymes with and without dockerin domains

Enzyme	**Without dockerin**			**With dockerin**		
	K_M[a]	V_{max} (U/mg)	V_{max} (U/mmol)	K_M[a]	V_{max} (U/mg)	V_{max} (U/mmol)
Xylanase, family 10 (59/70.6 kDa)[b]	3.3 ± 0.4	2.9 (± 0.5) X 10^5	4.9	3.1 ± 0.5	2.5 (± 0.4) x 10^5	3.5
Xylanase, family 11 (20.4/30.1 kDa)	13.2 ± 1.0	534 ± 81	2.6 x 10^{-2}	16.5 ± 1.2	370 ± 67	1.2 x 10^{-2}
Ferulic acid esterase (31.7/39.4 kDa)	0.77 ± 0.10	14.77 ± 1.8	3.2 x 10^{-4}	0.76 ± 0.15	4.10 ± 0.7	1.0 x 10^{-4}
Bifunctional ABF (59/68 kDa)	1.53 ± 0.2	120 ± 14	2.0 x 10^{-3}	1.75 ± 0.3	48 ± 8	7.0 x 10^{-4}

[a] mg/mL for Xylanase family 10 and 11, mM for ferulic acid esterase and ABF.
[b] Molecular weight of protein without dockerin / Molecular weight of protein with dockerin

3.3.1.2 Temperature effects on cohesin-dockerin interaction

Affinity constants for both *C. thermocellum* and *C. cellulovorans* cohesin/dockerin systems were calculated at multiple temperatures to assess the stability of xylanosome structures during hydrolysis reactions. An explanation of how SPR can aid in determining such values is discussed in Section 2.5. Table 3.2 shows the association (k_a or k_{on}) and dissociation (k_d or k_{off}) constants for each cohesin/dockerin system were on the same order of magnitude of reported K_A values for *C. thermocellum*, *Clostridium josui*, and *Clostridium cellulolyticum* at 25°C. The values slightly vary between enzymes with the same dockerin domain, which is either due to different molecular weight or native protein cellular location (cellulosomal versus non-cellulosomal) of each catalytic domain.

Table 3.2 Affinity constants for cohesin/dockerin domains at 25°C

Enzyme (*Cohesin/Dockerin System*)	**Literature Value (same species)**	**Reference**	**This Study**
X10Doc (*C. thermocellum*)	$< 10^9$ M^{-1}	[7]	4.4 (± 0.9) x 10^9 M^{-1}
X11Doc (*C. thermocellum*)	$< 10^9$ M^{-1}	[7]	5.8 (± 1.2) x 10^9 M^{-1}
FAEDoc (*C. cellulovorans*)	n/a		4.2 (± 0.8) x 10^8 M^{-1}
ABFDoc (*C. cellulovorans*)	n/a		3.9 (± 0.8) x 10^8 M^{-1}
C. cellulolyticum	$10^7 – 10^{10}$ M^{-1}	[16, 17]	n/a
C. josui	$10^9 – 10^{10}$ M^{-1}	[7]	n/a

Since xylan hydrolysis takes place at higher temperatures, cohesin/dockerin systems were analyzed at a range of temperatures from 25 – 60°C to determine the stability of xylanosomes during the reactions. As shown in Table 3.3 and Figure 3.1, increasing the temperature from 25°C to 60°C strengthened the overall interaction between cohesins and dockerins. However, the *C. thermocellum* system continued to have a higher affinity constant than *C. cellulovorans*. This may be related to the preferred temperatures of the gene origins. The *C. thermocellum* system is from a thermophillic bacterium, cultured at 60°C, while the *C. cellulovorans* system is mesophilic, suggesting that the cohesin/dockerin systems are optimized for their particular growth environment and adds another dimension to their species-specificity.

Table 3.3 Temperature effects on cohesin/dockerin domain affinity constants

Temperature	Cohesin/Dockerin System, K_A					
	C. thermocellum (X10Doc)			*C. cellulovorans* (FAEDoc)		
	k_{on} ($M^{-1}s^{-1}$)	k_{off} (s^{-1})	K_A (M^{-1})	k_{on} ($M^{-1}s^{-1}$)	k_{off} (s^{-1})	K_A (M^{-1})
25°C	9.1×10^{4}	2.1×10^{-5}	4.4×10^{9}	1.2×10^{5}	2.9×10^{-4}	4.2×10^{8}
40°C	3.0×10^{5}	4.7×10^{-5}	6.4×10^{9}	2.3×10^{5}	1.3×10^{-4}	1.7×10^{9}
50°C	6.7×10^{5}	7.2×10^{-5}	9.3×10^{9}	2.8×10^{4}	1.2×10^{-5}	2.4×10^{9}
60°C	6.9×10^{5}	6.6×10^{-5}	1.1×10^{10}	2.4×10^{5}	5.3×10^{-5}	4.4×10^{9}

Results are the average of two duplicate experiments with 20 – 25% error.

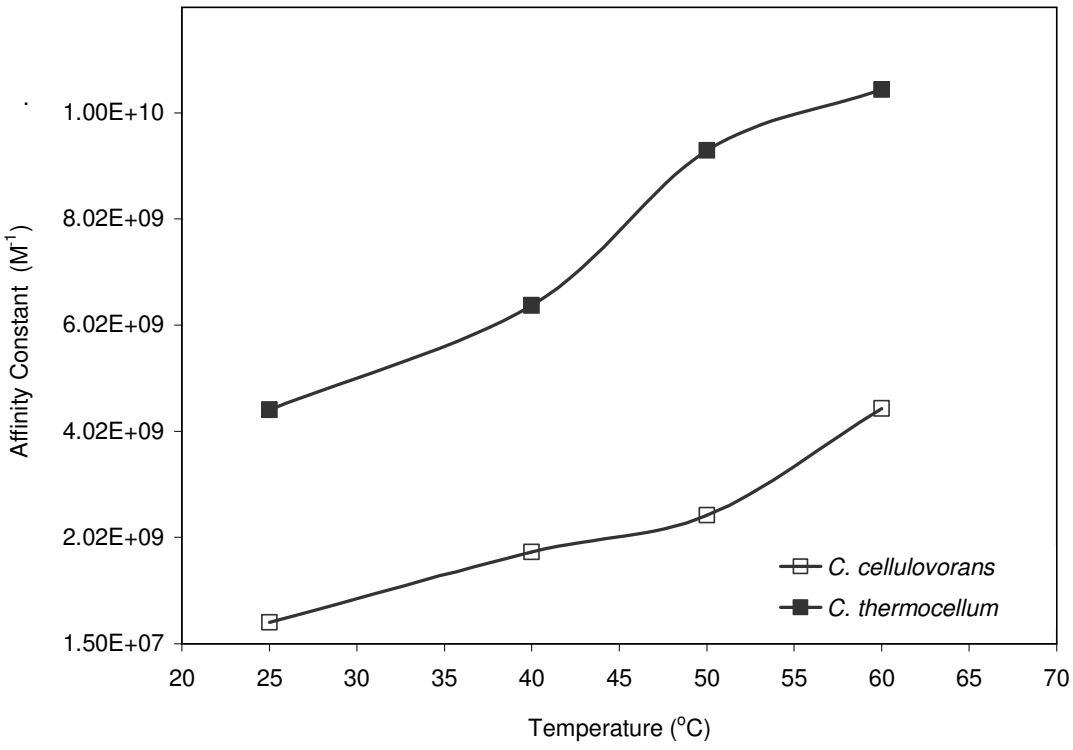


Figure 3.1 Temperature effects on cohesin/dockerin domain affinity constants.

With the ability to determine the dissociation constants, the enthalpies for the cohesin/dockerin interaction were estimated. Figure 3.2 shows a van't Hoff plot for each system. From the slopes obtained from linear regression of the data, the enthalpies of association were calculated as -13.01 kcal/mol for the *C. cellulovorans* system, and -5.09 kcal/mol for the *C. thermocellum* system, making the binding of dockerins with cohesin a weakly exothermic interaction. These values correlate well with those reported in literature for *C. thermocellum* (-2.22 kcal/mol), generated using isothermal titration calorimetry [18]. The ability of the cohesins and dockerins to maintain a high affinity interaction over a range of temperatures provides evidence for utilizing xylanosomes, as well as other protein nanostructures bearing these domains, at higher hydrolysis temperatures, as well as in consolidated bioprocessing applications.

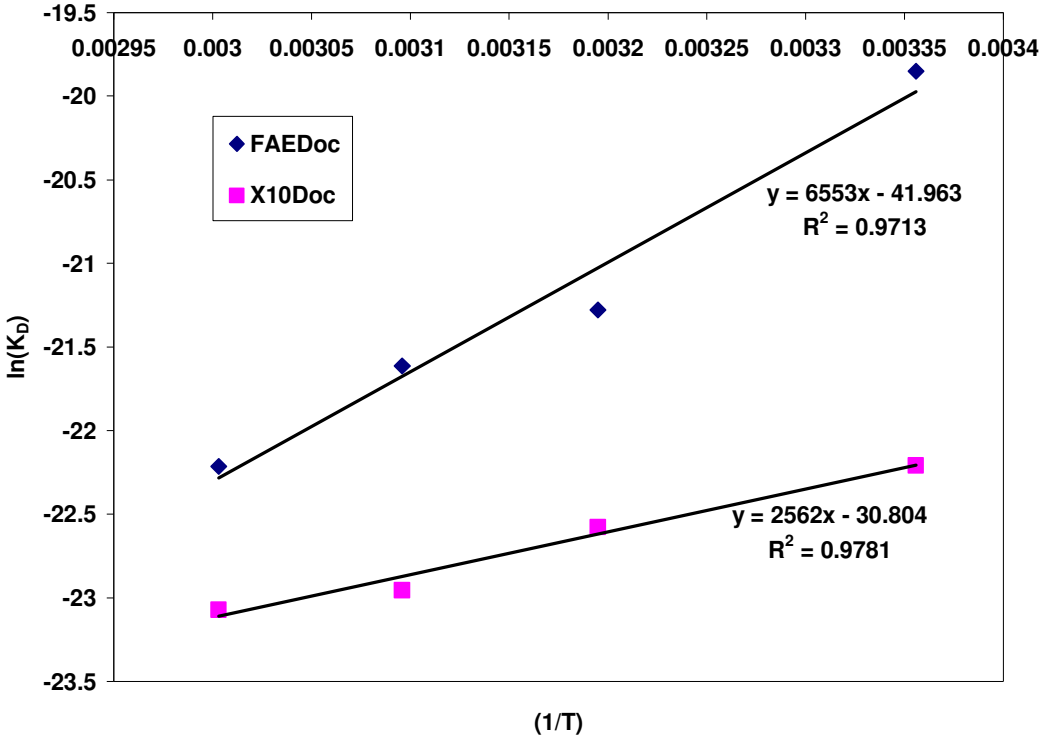


Figure 3.2 van't Hoff plot for binding of dockerin to cohesin domains for *C. thermocellum* and *C. cellulovorans*. Equations obtained using linear regression of the data.

3.3.1.3 Effects of "dirty" enzyme solutions on cohesin-dockerin interaction

Native cellulosomes assemble extracellularly, indicating that cohesins and dockerins successfully interact in the presence of non-specific proteins. However, the typical extracellular fraction of bacteria has a lower protein concentration than those usually found in crude lysates. In order to make use of self-assembling nanostructures in industrial applications, it would be ideal if cohesins and dockerins could maintain their high-affinity interaction from crude lysates, eliminating the need for protein purification prior to any enzymatic hydrolysis reactions. Affinity constants for *C. thermocellum* and *C. cellulovorans* systems were measured in the presence of significant amount of non-

dockerin tagged enzymes. Table 3.4 compares the rates of association and dissociation for 1% (w/v) and 20% (w/v) dockerin-tagged X10Doc.

Table 3.4 Affinity constants for *C. thermocellum* system in the presence of non-specific proteins

Analyte Concentration	**Rate of association (k_a, $M^{-1}s^{-1}$)**	**Rate of dissociation (k_d, s^{-1})**	**Affinity Constant (K_A, M^{-1})**
1% (w/v)	4.0×10^{6}	4.8×10^{-3}	8.4×10^{8}
20% (w/v)	7.1×10^{5}	2.8×10^{-3}	2.5×10^{8}
Purified	9.1×10^{4}	2.1×10^{-5}	4.4×10^{9}

Results are the average of two duplicate experiments with 15 – 25% error.

The selected concentrations were meant to simulate both low (1%) and high (20%) levels of recombinant protein expression in *E. coli*. As listed in Table 3.4, the affinity constants are an order of magnitude lower in the presence of 99% non-specific protein (crude lysate from X11-without dockerin), compared to a K_A of 4.4×10^{-9} M^{-1} for the purified protein at the same conditions. Furthermore, rates of dissociation were comparable for 1% and 20% solutions, but about two orders of magnitude faster than the purified solution. The rates of association increased with lower percentages of X10Doc, which may be an artifact of more non-specific interactions with the scaffolding protein in addition to the specific cohesin/dockerin interaction. In general, xylanosomes can assemble from 'crude' lysates, providing yet another advantage in using these structures industrially.

3.3.1.4 Effects of ethanol on cohesin-dockerin interaction

Consolidated bioprocessing is valuable in the simultaneous saccharification and fermentation of sugars from biomass to ethanol. The use of xylanosomes within these processes requires an understanding of how these structures can form in the presence of ethanol, providing another attribute of these cellulosome-like nanostructures. Ethanol was added to 10% (v/v) to the running buffer to measure the affinity constants of both *C. cellulovorans* and *C. thermocellum* cohesin-dockerin systems using SPR maintained at 40^{o}C. This concentration was selected as a maximum value obtained from fermentations with engineered microbes [19]; the temperature represents that near to *E. coli* microbial fermentations. Table 3.5 lists the K_A values with and without the presence of ethanol.

Table 3.5 Affinity constants for *C. thermocellum* and *C. cellulovorans* in the presence of ethanol

Condition	Cohesin/Dockerin System, K_A					
	C. thermocellum (X10Doc)			*C. cellulovorans* (FAEDoc)		
	k_{on} ($M^{-1}s^{-1}$)	k_{off} (s^{-1})	K_A (M^{-1})	k_{on} ($M^{-1}s^{-1}$)	k_{off} (s^{-1})	K_A (M^{-1})
0% (v/v) EtOH, 40°C	3.0×10^5	4.7×10^{-5}	6.4×10^9	2.3×10^5	1.3×10^{-4}	1.7×10^9
10% (v/v) EtOH, 40°C	7.4×10^4	7.5×10^{-4}	9.8×10^7	1.4×10^4	1.0×10^{-4}	1.4×10^7

Results are the average of two duplicate experiments with 20-25% error.

Overall, ethanol lowers the affinity constants two orders of magnitude compared to those measured without ethanol. Both the rate of binding and dissociation of dockerins from the immobilized scaffolding are significantly reduced, indicating that ethanol inhibits the proper interaction of the two domains to form the protein structure. The cohesin/dockerin interaction has been reported as hydrophobic [18]. Ethanol in the buffer provides less stringent solution conditions by allowing changes in water-water and water-ethanol molecular interactions. Thus, the ethanol-water solution is more hydrophobic, allowing the critical residues in the cohesin-dockerin binding mechanism to interact with the solution more frequently than with each other compared to solutions without ethanol. However, it is important to note that the affinity constants are still relatively strong, allowing for construction of xylanosome or designer cellulosomes in the presence of ethanol and thus providing another positive attribute to these structures for use in consolidated bioprocessing.

3.3.2 Biomass hydrolysis using two-unit xylanosomes

3.3.2.1 Release of reducing sugars from wheat arabinoxylan

Wheat arabinoxylan (WAX) was used to test the efficiency of the two-unit xylanosomes on xylan hydrolysis. All four xylanosome combinations and corresponding free enzyme mixtures were applied to 0.25, 0.50 and 1.00% (w/v) WAX and the concentration of reducing sugars analyzed after 24 hours. Results demonstrate the ability of xylanosomes to improve the release of reducing sugars. Table 3.6 reveals that the most effective xylanosome contains X10Doc with ABFDoc at 1% (w/v) giving a 31% increase in the concentration of released reducing sugars, followed by X11Doc and ABFDoc, at 0.5% (w/v) then 1% (w/v) WAX with 21-22% increase in reducing sugars after 24 hours of hydrolysis.

Table 3.6 Structure-endowed synergy observed in the release of soluble reducing sugars from various concentrations (% w/v) of wheat arabinoxylan [c]

Enzyme Combination	**Synergy**					
	6 hours			24 hours		
	0.25%	*0.50%*	*1.00%*	*0.25%*	*0.50%*	*1.00%*
X10Doc/ FAEDoc	0.95	1.05	1.01	0.93	1.03	0.96
X11Doc/ FAEDoc	1.02	1.10	1.05	1.03	1.06	1.01
X10Doc/ ABFDoc	0.94	0.96	1.26	0.91	0.93	1.31
X11Doc/ ABFDoc	0.99	1.08	1.19	0.97	1.22	1.15

[c] Synergy is calculated by dividing the concentration of product released by enzyme mixtures by the sum of the concentrations of product released by each individual enzyme. Calculated synergies are the average of three independent hydrolysis experiments.

This synergy is anticipated since the ABFDoc enzyme can remove arabinose side chains, as well as cleave single xylose molecules from xylo-oligosaccharides resulting from xylanase activity; both activities result in the release of more reducing sugars. Conversely, none of the xylanosomes containing FAEDoc significantly improved the release of reducing sugars, as expected, since only the xylanase would contribute to their accumulation. Additionally, the concentration of biomass selected can affect the performance of xylanosomes. For example, an ideal concentration of 1% (w/v) WAX gives synergy for ABFDoc xylanosomes, whereas a lower concentration of 0.25% (w/v) never provides conditions for synergistic release of sugars to occur with xylanosomes.

Figure 3.3 displays how the use of xylanosomes, as well as the choice of xylanase employed in the xylanosome, improves release of sugars over time. After six hours, ABFDoc xylanosome reactions already show improvement over the free enzyme systems, with this trend continuing after 24 hours of incubation. Furthermore, structures and free enzyme mixtures containing X10Doc resulted in higher concentrations of reducing sugars, 5.5 mM and 4.2 mM, respectively, compared to those with X11Doc (between 3.2 – 3.9 mM). This trend connects to the preference of each xylanase and the complexity of WAX: family 10 xylanases prefer highly-substituted xylan backbones, whereas family 11 xylanases prefer less-substituted regions [20]. X10Doc has greater activity than X11Doc on WAX, as observed in the free enzyme systems. Moreover, the presence of a family 22 CBM targets xylan and most likely positions the X10Doc enzyme and/or xylanosome structure to be more effective at cleaving the xylan backbone. With the distinct goal of increasing the amount of reducing sugars in the hydrolysate,

xylanosomes containing a xylanase and bi-functional arabinofuranosidase / xylosidase are more effective at than their corresponding free enzyme systems.

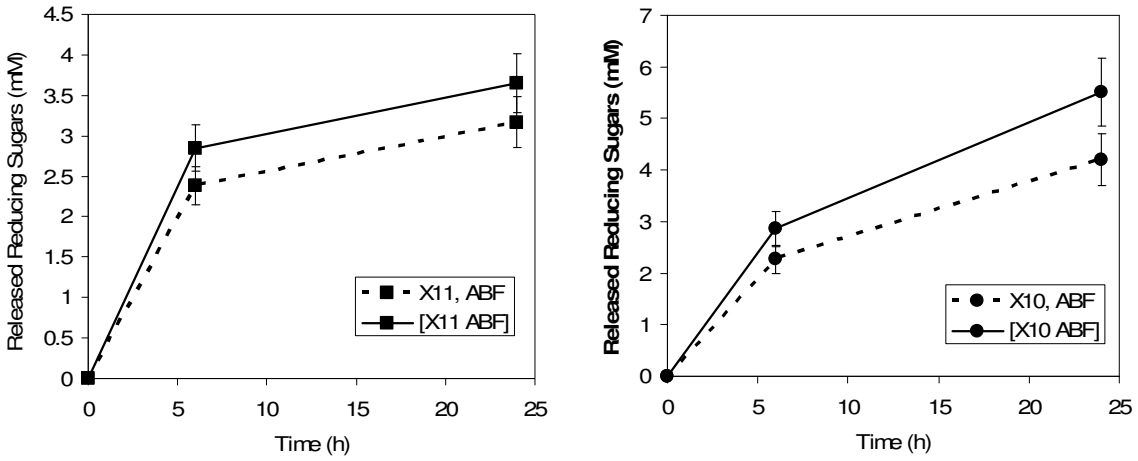


Figure 3.3 Release of reducing sugars from wheat arabinoxylan. Results are from measurements of three independent experiments. Error bars represent standard deviation.

3.3.2.2 Release of ferulic acid from wheat arabinoxylan

In addition to fermentable sugars for bioethanol production, ferulic acid is another valuable chemical available from hemicellulose, as it can be used in the bioproduction of vanillin [21]. Work described in Chapter 4 reveals synergistic release of ferulic acid when either a xylanase from *C. thermocellum* and *Bacillus halodurans* was combined with a ferulic acid esterase from *Cellvibrio japonicus* from corn bran but not wheat arabinoxylan. Figure 3.4 reveals use of FAEDoc-containing xylanosomes instead of free enzymes increases the amount of ferulic acid removed from WAX. Table 3.7 shows the most synergistic xylanosomes contain X10Doc and FAEDoc applied to 0.5% (w/v) WAX, followed by X11Doc and FAEDoc at the same conditions, releasing 14 and 18

μM ferulic acid, respectively. X10Doc and FAEDoc xylanosomes were also effective at 1% (w/v) WAX. Reactions with enzyme mixtures containing ABFDoc failed to release ferulic acid into solution. Furthermore, it appears that the use of a lower concentration of WAX (0.25% w/v) is overall not beneficial in the use of xylanosomes, but 0.5% (w/v) is an optimal concentration in the release of ferulic acid.

Table 3.7 Structure-endowed synergy observed with two-unit xylanosomes for the release of ferulic acid from various concentrations (% w/v) of wheat arabinoxylan [d]

Enzyme Combination	**Synergy**					
	6 hours			24 hours		
	0.25%	*0.50%*	*1.00%*	*0.25%*	*0.50%*	*1.00%*
X10/FAE	0.94	1.17	1.06	0.88	1.21	1.13
X11/FAE	0.93	1.03	1.07	1.10	1.13	1.08

[d] Synergy is calculated by dividing the concentration of product released by enzyme mixtures by the sum of the concentrations of product released by each individual enzyme. Calculated synergies are the average of three independent hydrolysis experiments.

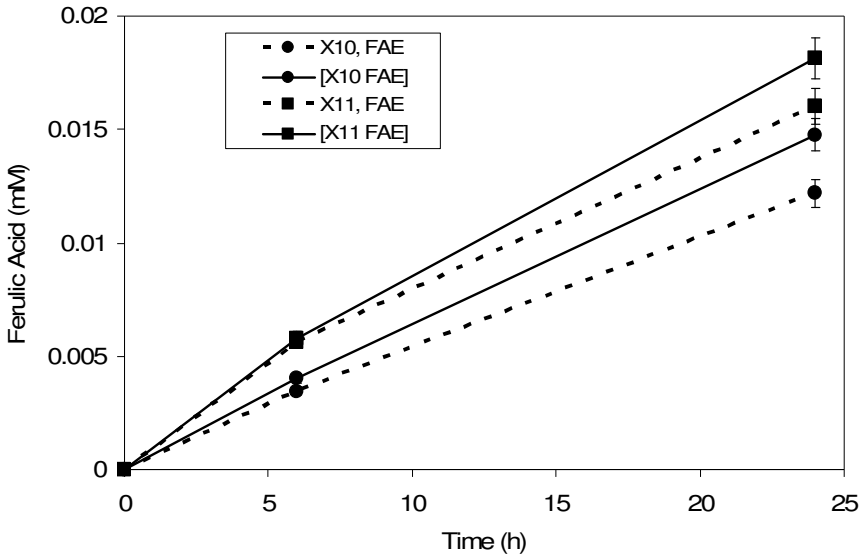


Figure 3.4 Release of ferulic acid from wheat arabinoxylan. Results are from measurements of three independent experiments. Error bars represent standard deviation.

Interestingly, Figure 3.4 shows xylanosomes and free enzyme mixtures containing X11Doc performed better at the release of ferulic acid than those containing X10Doc, which is contrary to results relating to the release of reducing sugars. Reasoning behind this trend leads back to the preference of each xylanase; perhaps the action of FAEDoc provides more regions of less substituted xylan that X11Doc can cleave and reduces the amount of preferred substrate for X10Doc.

These results show that combining a xylanase and ferulic acid esterase in a protein nanostructure gives a synergistic release of ferulic acid compared to free enzyme mixtures and xylanase selection can lead to increased concentrations of ferulic acid from WAX. In general, xylanosomes can be used to improve release of reducing sugars or ferulic acid by varying the enzymatic makeup of the structure.

3.3.2.3 Release of reducing sugars and ferulic acid from destarched corn bran

Xylanosomes were combined with two cellulases found within cellulosomes of *C. thermocellum*: Cel9A, a family 9 glycosyl hydrolase (GH) with endoglucanase activity, and Cel48S, a family 48 glycosyl hydrolase possessing cellobiohydrolase activity. Destarched corn bran (DCB) was used to test the ability of xylanosomes to increase access to cellulose compared to free, unstructured systems. Figures 3.5 and 3.6 show improvement in the release of reducing sugars from DCB using structured over free hemicellulases in conjunction with Cel9A or a combination of Cel9A and Cel48A. Table 3.8 lists the resulting synergies observed during the hydrolysis reactions.

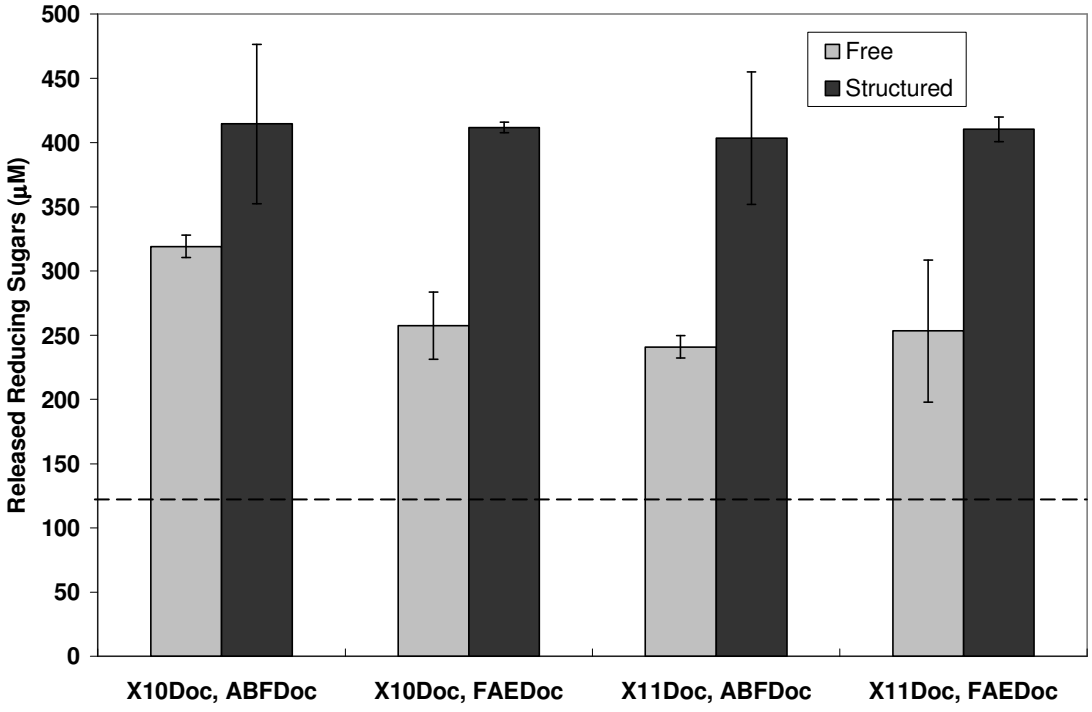


Figure 3.5 Release of reducing sugars from destarched corn bran with free and structured hemicellulases in conjunction with Cel9A. Cel9A alone (dotted line) released 122 ± 35 μM reducing sugars. Results are from measurements of three independent experiments. Error bars represent standard deviation.

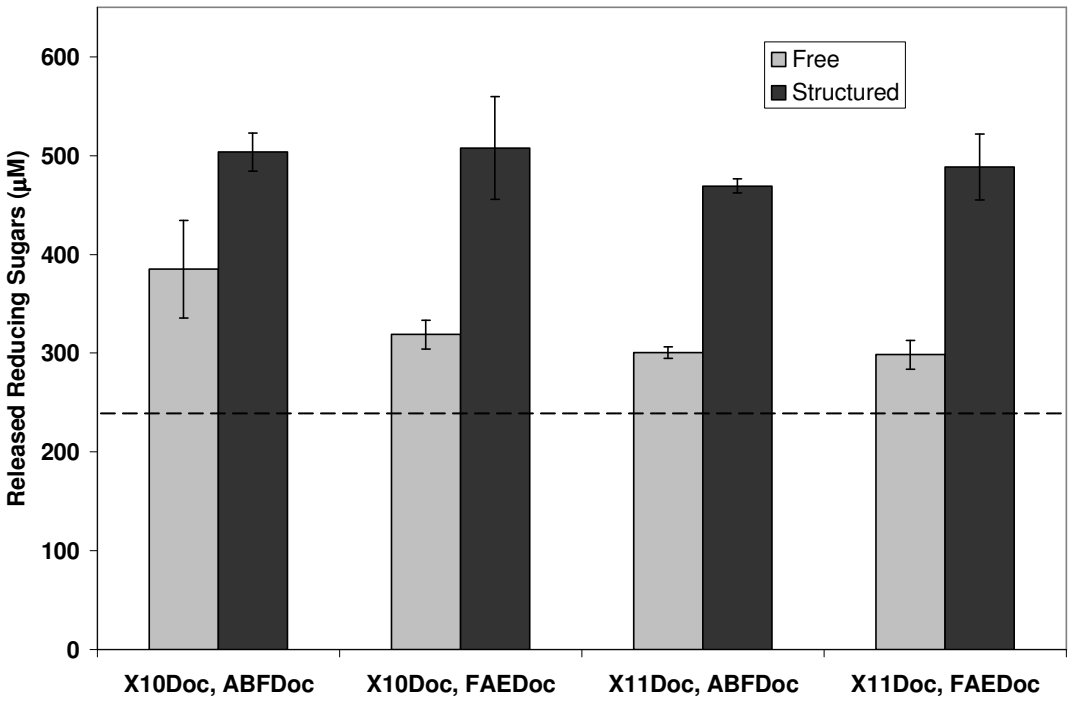


Figure 3.6 Release of reducing sugars from destarched corn bran with free and structured hemicellulases in conjunction with Cel9A plus Cel48A. Cel9A plus Cel48A (dotted line) released 235 ± 12 μM reducing sugars. Results are from measurements of three independent experiments. Error bars represent standard deviation.

From Figures 3.5 and 3.6, addition of both free and structured hemicellulases improves the release of reducing sugars. Although the selected hemicellulases do not release a detectible amount of reducing sugars themselves in either form, xylanosomes allow CelA and CelS to be more active on DCB than just adding free hemicellulases to the reaction. Compared to the sugars released by cellulases alone, it appears that xylanosomes are more beneficial in increasing the accessibility to the cellulose present in destarched corn bran (DCB) than free hemicellulases. The structure-endowed synergies are listed in Table 3.8.

Table 3.8 Synergistic effects of adding CelA and CelS with 2-unit xylanosomes on the release of sugars from destarched corn bran [e]

Enzyme Combination	Synergy	
	Xylanosomes + Cel9A	Xylanosomes + Cel9A + Cel48S
X10Doc/ABFDoc	1.30 (1.48)	1.31 (1.79)
X10Doc/FAEDoc	1.60 (2.14)	1.59 (3.24)
X11Doc/ABFDoc	1.68 (2.37)	1.56 (3.56)
X11Doc/FAEDoc	1.62 (2.20)	1.64 (3.98)

[e] Combining Cel9A and Cel48S gave a synergy of 1.53 compared to the sum of sugars released using individual enzymes. Neither free nor structured hemicellulases release detectible sugars from DCB. Synergy is calculated as described previously. Synergy values in parentheses are calculated similarly, except RRS concentrations from cellulases alone were subtracted from RRS concentrations released by cellulase/hemicellulase mixtures.

Table 3.8 shows that all xylanosomes are effective at improving the release of reducing sugars from DCB in conjunction with cellulases. When looking at the additional sugars released due to the presence of hemicellulases, use of structured enzymes is 1.5 to 2.4-fold more advantageous than free enzymes in the case of Cel9A mixtures. Structured enzymes show even greater effectiveness (1.8 to 4-fold) in the presence of both Cel9A and Cel48S cellulases, creating a more robust enzyme cocktail with the addition of the SP2 scaffolding protein. This structure-endowed synergy is most likely due to the close proximity of multiple hydrolyzing enzymes to the hemicellulose within the biomass substrate. A parallel is seen within the *C. thermocellum* cellulosome. The presence of bi-functional proteins with xylanase and ferulic acid esterase activities (*XynY/XynZ*) help cellulases gain access to naturally encountered complex substrates,

where cellulose is surrounded by hemicellulose [22]. Furthermore, free enzymes perhaps perform a more random cleavage of the hemicellulose polymer, whereas xylanosomes create larger areas (i.e., holes) within hemicellulose due to their structured form, providing greater accessibility to cellulose. By increasing accessibility, hydrolysis is improved.

On wheat arabinoxylan, two-unit xylanosomes containing FAEDoc provide some improvement in release of sugars and ferulic acid. However, on DCB, ferulic acid was not synergistically released with xylanosomes containing X10Doc, although xylanosomes with X11Doc did show slight improvements. As shown in Figure 3.7, all of the free enzyme systems remove similar concentrations of ferulic acid. Furthermore, structures with X10Doc and FAEDoc seem to inhibit ferulic acid release with either cellulase combination.

While the presence of the family-22 CBM may aid in sugar release, it may restrict xylanosome movement for the less frequent ester-linked ferulic acid moieties. Similarly, lower concentrations of ferulic acid were released from WAX with X10Doc-FAEDoc compared to X11Doc-FAEDoc xylanosomes. This inhibition is not seen with X11Doc xylanosomes, which do show modest synergy for ferulic acid removal from the xylan backbone. Based on data from both WAX and DCB, it seems that combining a family 11 xylanase with ferulic acid esterase performs better than a family 10 xylanase. Moreover, CBMs may impede access to ferulic acid, also resulting in lower ferulic acid release.

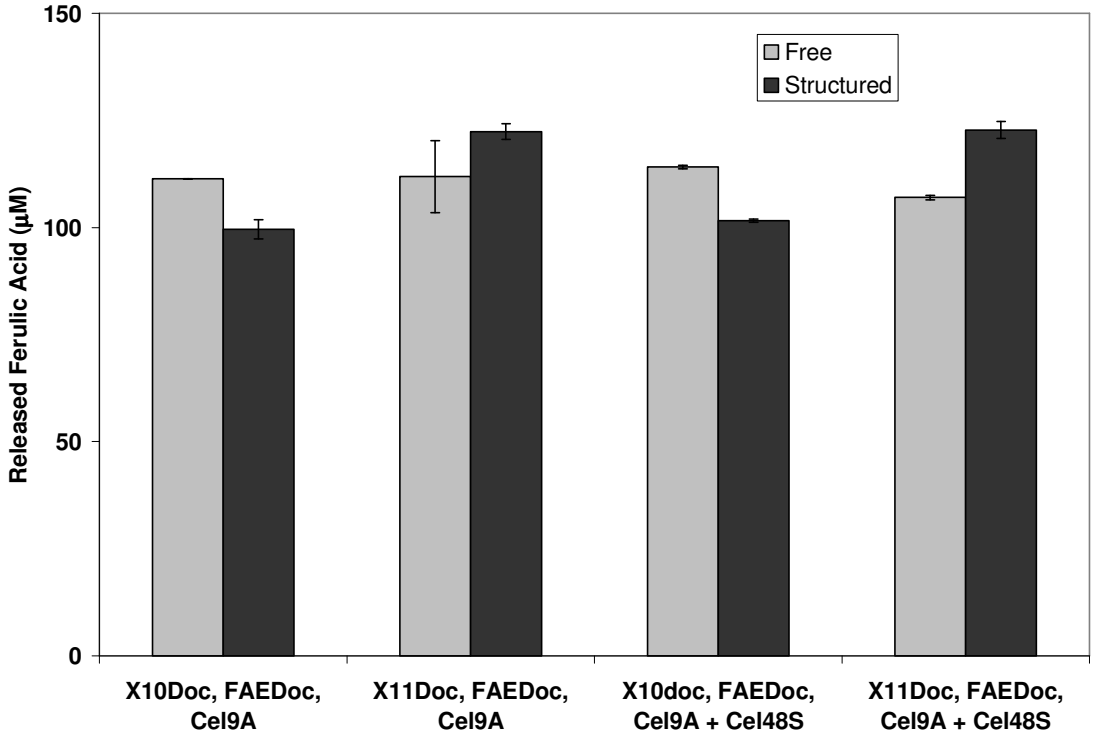


Figure 3.7 Release of ferulic acid from destarched corn bran with free and structured hemicellulases in conjunction with Cel9A and Cel48A. Cel9A, Cel48A, nor a mixture of the two released any ferulic acid from DCB. Results are from measurements of three independent experiments. Error bars represent standard deviation.

3.3.3 Cell-surface two-unit xylanosomes for possible consolidated bioprocessing

To display SP2 onto the cell-surface of *E. coli*, DNA sequences encoding the signal peptide from the outer membrane protein, *ompA*, followed by the first nine amino acids of the lipoprotein *lpp* to anchor scaffolding into the cell membrane, both from *E. coli,* were fused to the N-terminal of SP2 nucleotide sequence and inserted into pSTV28 vector. The recombinant vector was used to transform *E. coli* UT5600, resulting in a strain expressing SP2 on the cell outer membrane. After induction of SP2, cell pellets were collected and mixed with purified xylanosome enzymes in the presence of calcium for formation of the two-unit xylanosomes. These cells were used to perform whole-cell

reactions on synthetic substrates to confirm presence of constructed xylanosomes on the cell-surface of *E. coli*.

Figures 3.8 and 3.9 show that *E. coli* displaying the scaffolding protein on the surface retained activity on both methyl ferulate and oat spelt xylan when appropriate enzymes were individually added to form xylanosomes. When two enzymes were added to form a complete xylanosome, the activities toward methyl ferulate and oat spelt xylan decreased 80 percent and 25 percent, respectively. Lower activity suggests that enzymes within the scaffolding can hinder the access of target substrates of other xylanosome-bound enzymes. Another possible explanation is the cohesin/dockerin interaction is not completely species-specific. The addition of one dockerin-tagged enzyme (FAEDoc, X10Doc, or X11Doc only) allows for some level of cross-species interaction with cohesins present in the scaffolding, leading to more enzymes on the cell-surface and thus, higher activity. Despite reduced activity, constructing xylanosomes on the cell surface of *E. coli* opens the possibility of consolidated bioprocessing methods for xylan hydrolysis.

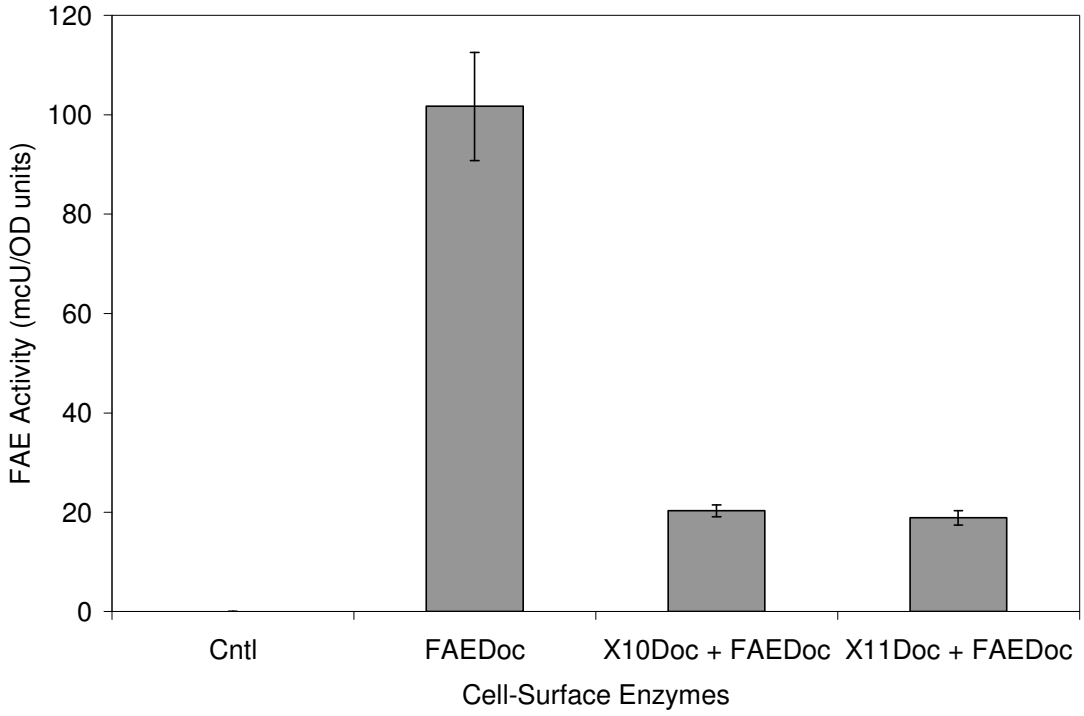


Figure 3.8 Whole-cell xylanosome ferulic acid esterase activities. Substrate: methyl ferulate. Results are from measurements of two independent experiments. Error bars represent standard deviation. Control sample was whole-cell reaction without expressed scaffolding protein.

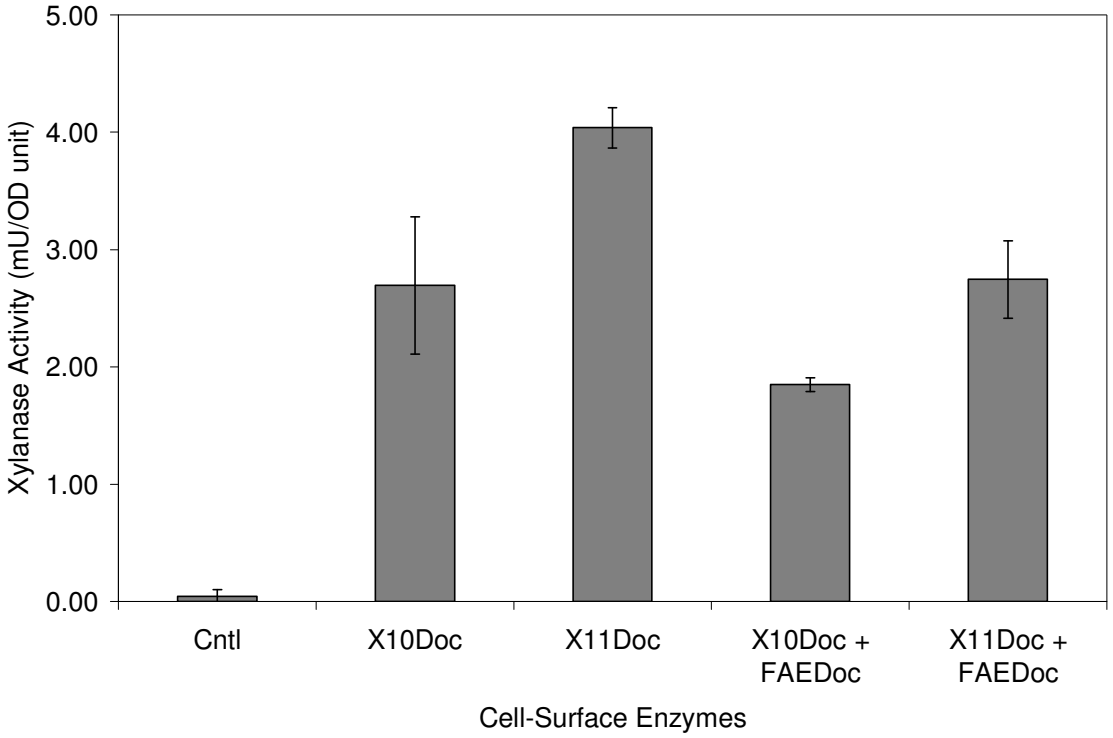


Figure 3.9 Whole-cell xylanosome xylanase activities. Substrate: Oat spelt xylan. Results are from measurements of two independent experiments. Error bars represent standard deviation. Control sample was whole-cell reaction without expressed scaffolding protein.

3.4 Discussion

Discovering a more efficient enzymatic method for xylan hydrolysis could be beneficial to lowering the cost of biomass utilization, especially in improving the overall sugar yield from biomass for bio-based chemical production. The multi-functional, self-assembling, highly-effective cellulosomes of anaerobic bacterial and fungi provide a platform from which to construct such a method. The architecture of cellulosomes, mainly cohesin and dockerin domains, was used to place hemicellulases in structured form using a chimeric scaffolding protein and dockerin-tagged enzymes. In this study, four different xylanosomes were constructed and tested for improved release of reducing

sugars and ferulic acid from wheat arabinoxylan (WAX) and destarched corn bran (DCB) and compared to free enzyme systems.

Each xylanosome contained two different hemicellulases: one backbone-acting enzyme, a xylanase (X10Doc, X11Doc), and one side-chain acting enzyme, either a ferulic acid esterase or a bi-functional arabinofuranosidase/xylosidase (FAEDoc, ABFDoc). On WAX, two-unit xylanosomes containing either xylanase and ABFDoc gave between 15 – 30% greater release of soluble reducing sugars from insoluble xylan than the equivalent enzyme mixture without the scaffolding protein. The preferred xylanase was from glycosyl hydrolase family 10, which contained a family 22 carbohydrate binding module (CBM). Previous work with designer cellulosomes showed that the presence of a CBM within the scaffolding protein gave higher synergistic release of sugars than those without a CBM [11]. Thus, close proximity of the enzyme (or xylanosome) to the substrate aids in better hydrolysis in addition to the presence of multiple hydrolases. Furthermore, the preferred substrate of each xylanase could also play a significant role in sugar release. Since the arabinoxylan backbone is likely to be highly substituted, X10Doc would naturally have a higher activity than X11Doc. Similarly, two-unit xylanosomes containing either xylanase and the FAEDoc improved the release of ferulic acid from the xylan backbone of WAX over the corresponding unstructured, free enzymes by 17 – 21%. In this case, combinations with X11Doc showed greater ferulic acid release, whereas X10Doc mixtures gave higher synergies. With DCB, X10Doc xylanosomes also gave lower concentrations of ferulic acid than their X11Doc counterparts. While it is unclear if the different substrate preference of X10Doc or the presence of a CBM is the cause of lower ferulic acid release, it is evident

that X11Doc with FAEDoc is a better enzyme mixture. Lastly, these xylanosomes did not show synergy in the release of sugars, being that only one enzyme could produce soluble sugars.

Native cellulosomes have been found to possess a combination of cellulases and hemicellulases in the presence of different biomass substrates, allowing for removal of hemicellulose to expose the target cellulose polymer [23, 24]. This information lead us to investigate whether xylanosomes play a better role by exposing cellulose of a more complex substrate containing both cellulose and hemicellulose than just xylan hydrolysis. Xylanosomes combined with either Cel9A (GH family 9, endoglucanase) or a combination of Cel9A and Cel48S (GH family 48, cellobiohydrolase) from *C. thermocellum* showed improvement on the release of sugars from destarched corn bran compared to the corresponding free enzyme systems. Calculated synergies on DCB were up to 60% higher than those on WAX. Xylanosomes were found to improve the release of sugars on DCB up to four times more than free enzyme systems with the addition of cellulases, creating a more effective enzyme cocktail. Furthermore, select xylanosomes removed more ferulic acid from DCB than free enzyme. Thus, xylanosomes could be added to enhance the enzymatic hydrolysis of a variety of biomass substrates beyond those that are rich in hemicellulose.

Successful construction of three-unit mini-cellulosomes on the cell surface of *Saccharomyces cerevisiae* has led to simultaneous saccharification and fermentation of cellulose to ethanol [25, 26]. This work displayed the possibility of developing a xylanolytic *E. coli* strain which utilizes xylanosomes for saccharification of xylan. Formation of xylanosomes on the cell surface was confirmed via activity on synthetic

substrates, however not on a complex biomass substrate, such as WAX or DCB. Testing of whole-cell xylanosomes on these substrates would be difficult, as enzymatic activity would release sugars for cell growth and skew overall hydrolysis results, and is beyond the scope of this dissertation. Nonetheless, analysis of the affinity constants of cohesin/dockerin systems from two *Clostridia* species reveal that cellulosome-like protein nanostructures possess positive attributes making them advantageous to use in consolidated bioprocessing (CBP). Cohesins and dockerins maintain high-affinity interactions at a range of temperatures applied in CBP, as well as in the presence of ethanol. Furthermore, these nanostructures can form from crude lysates, eliminating the need to purify components prior to in-situ construction, providing a strong basis for inclusion in CBP developments. Further work on developing a xylanolytic xylanosome-producing *E. coli* that can hydrolyze and subsequently ferment xylan into target products, such as biofuel or vanillin, would further provide evidence for consolidated bioprocessing with xylanosomes.

While xylanosomes show improvements in hydrolysis of some biomass, designer cellulosomes can produce synergies up to seven times that of matching free enzyme systems. Our synergies are on the lower end of this continuum, which may be a result of the presence of only two enzymatic activities within the xylanosome. Six different enzymatic activities are required for complete hydrolysis of xylan and perhaps incorporating more than two could result in higher synergies and is worth investigation. Next steps are to develop a three-unit xylanosome with a scaffolding protein containing three divergent cohesins incorporating a xylanase via X10Doc or X11Doc, an arabinofuranosidase and xylosidase by ABFDoc and a ferulic acid esterase and acetyl

xylan esterase from a newly characterized Fee1B from *Cellvibrio japonicus*., which will be discussed in Chapter 4. These three different enzymes provide a combination of five out of six total activities required to break down xylan into its monomeric components.

3.5 Materials and Methods

3.5.1 Strains and materials

Strains containing the dockerin-tagged enzymes (X10Doc, X11Doc, FAEDoc, and ABFDoc) and two-cohesin scaffolding protein (SP2) and expression plasmids were described in the materials and methods section in Chapter 2. *E. coli* UT5600/pBBTH-HHHR for cell-surface display of SP2, *E. coli* E609Y/pQpelCelA for leaky expression of Cel9A from *C. thermocellum*, and *E. coli* E609Y/pQTHCelS for leaky expression of Cel48S from *C. thermocellum* were constructed by and obtained from Dr. Hyun-Dong Shin. Purified deAX (ABF) protein without dockerin domain was graciously provided by Dr. Kurt Wagschal from USDA. Wheat arabinoxylan (Megazyme, Ireland) and destarched corn bran (SunOpta, Bedford, MA, U.S.A.) were complex substrates used application of xylanosomes and free enzymes in the release of reducing sugars and ferulic acid. Destarched corn bran was prepared according to Wang [27] using Termayl (Novozymes, Denmark) instead of G-ZYME.

3.5.2 Cloning and expression of enzymes without dockerin domain

The GH family 10 xylanase domain and family 22 carbohydrate binding domain from the *XynC* gene was amplified from the genomic DNA of *C. thermocellum* using the primers XynC nod-F: 5'—CGATAGGATCCGCAGCTCTGATTTACGATGATTT—3' and XynC nod-R: 5'—TATTGCTCGAGAACTATAGCATAAAATGCAGGTTTTG—

3' with *Bam*HI and *Xho*I restriction sites (underlined), respectively. The resulting PCR fragment was digested and inserted into *Bam*HI-*Xho*I linearized pET20b(+) and transformed into *E. coli* BL21Star (DE3) for expression. Similarly, the ferulic acid esterase domain from *XynZ* was isolated from the genome of *C. thermocellum* using primers FAE-F and FAE-R listed in Table 2.1, inserted into *Bam*HI-*Sal*I linearized pET20b(+), and transformed into *E. coli* BL21Star (DE3) for expression. The GH family 11 xylanase was cloned using Xyl11A-F and Xyl11A-R primers listed in Table 2.1 and inserted into *Bam*HI-*Hind*III linearized pQE80L for expression using *E. coli* JM109. All three enzymes were purified via immobilized metal affinity chromatography using the protocol described in the materials and methods section of Chapter 2, Section 2.3.

3.5.3 Calculation of Michaelis-Menten constants, K_M and V_{max}

Enzyme assays were performed as described in Section 2.3.8 using a range of concentrations of substrates: oat spelt xylan (0.1% (w/v) – 2.5% (w/v)), methyl ferulate (0.1 mM – 5 mM), and *p*-nitrophenyl β-D-xylopyranoside (0.025 mM – 4 mM). Assays were repeated in duplicate and K_M and V_{max} values were calculated from the slope and y-intercept, respectively, of a linear regression on a double reciprocal plot (Lineweaver-Burke) of substrate concentration versus enzyme activity for each enzyme with and without a dockerin domain.

3.5.4 Surface plasmon resonance

Immobilization of SP2 was performed as described in Section 2.3.9. Running buffer was 20 mM Tris-maleate, 1 mM CaCl2, 0.05% (v/v) Tween 20, pH 6.0 with or without 10% (w/v) ethanol (200 proof, Acros Organics, U.S.A), depending on the nature of the experiment. Dockerin-tagged enzymes were exchanged into the Tris-maleate

running buffer and diluted to between 10 and 100 nM concentrations and allowed to interact with the immobilized SP2 for 180-s injections to measure rate of association. Running buffer was flowed over the surface immediately following injections for 120 seconds to measure rate of dissociation. The slide surface was regenerated by removing bound dockerin-tagged enzymes using an injection of 15 mM HCl for 90 to 120 seconds. Three different concentrations of analyte were injected in duplicate for determining binding kinetics. In the case of temperature studies, the system was equilibrated to the designated temperature for at least 15 minutes prior to running injections. Various concentrations of analytes were generated by combining the non-binding fraction from the purification of a non-dockerin tagged enzyme (Xyl11A) with purified X10Doc. Resulting spectra were evaluated using the Scrubber analysis software (Reichert) to measure the affinity constants of both cohesin-dockerin pairs. Data were interpreted using a first-order binding model, where the dockerin-tagged ligand, L, plus the immobilized acceptor, A, form a complex, LA, in a 1:1 ratio. As a control, enzymes were also injected over the gold slide without the immobilized SP2. Initial exposure or injections of known concentrations of analyte over the immobilized protein provides the rate of association, k_{on}, for the analyte-protein complex and flowing running buffer in the absence of analyte determines the rate of dissociation, k_{off}, of the analyte from the immobilized protein. From these values, an affinity constant, K_A, defined as k_{on}/k_{off}, or dissociation constant, $K_D = k_{off}/k_{on}$, is determined.

3.5.5 Xylanosome construction and biomass hydrolysis reactions

Protein concentration of xylanosome components was measured by the method described by Bradford using a protein reagent dye (Bio-Rad), with bovine serine albumin

as the standard. Xylanosome construction was performed by adding equimolar quantities of appropriate dockerin-tagged enzymes in 20 mM Tris-maleate, 1 mM $CaCl_2$, pH 6.0 at room temperature in 1.5-mL microcentrifuge tubes. After brief centrifugation, equimolar quantities of the scaffolding protein, SP2, were added to the solution for structure formation. For xylanosomes containing ABFDoc, ABFDoc was combined with SP2 first, then either X10Doc or X11Doc added, to complete the xylanosome. Proteins were added to obtain a final concentration of 0.1 μM for a 1 mL reaction volume.

E. coli E609Y/pQpelCelA and *E. coli* E609Y/pQTHCelS were cultured in LB with ampicillin (100 μg/mL) and induced with 0.5 mM IPTG for 20 hours at 22°C for extracellular expression of Cel9A and Cel48S. Cells were removed via centrifugation (4000 x *g*, 25 minutes, 4°C). Resulting supernatants were concentrated three times using dialysis tubing (MW 10,000 Da cutoff) and polyethylene glycol (Sigma-Aldrich). Cellulase activity was measured using 1% (w/v) carboxymethyl cellulose (CMC, 30 minutes, 40°C) and reducing sugars determined as described by Miller (1959). Cel9A activity was 1024 U/mL concentrated supernatant. Cel48S had no detectible activity on CMC.

For hydrolysis of biomass substrates, 5% (w/v) wheat arabinoxylan or 6% (w/v) destarched corn bran in 20 mM Tris-maleate, 1 mM CaCl2, pH 6.0, were added to xylanosomes and incubated at 50°C for 24 hours at 1000 rpm in a thermomixer. In reactions with cellulases, 150 U (or 150 μL) of Cel9A and 150 μL of Cel48S were added to the reaction. To stop the hydrolysis reaction, mixtures were exposed to boiling water for 5 minutes. After cooling, reaction mixtures were centrifuged for 10 minutes at 16000 x *g* to pellet insoluble biomass and the supernatant removed for subsequent analysis.

To measure released reducing sugars, appropriately diluted supernatants were combined with dinitrosalicyclic (DNS) acid solution as outlined in Miller [28] and incubated in boiling water for 5 minutes for the colorimetric reaction to occur. After cooling, the absorbance at 550 nm was measured and compared to a calibration curve generated with xylose as the standard. Quantification of released ferulic acid was performed on hydrolysis supernatants by high performance liquid chromatography as described by Shin and Chen [29].

3.5.6 Cell-surface xylanosome reactions

Cell-surface expression of SP2 protein on the surface of *E. coli* was performed as follows. Luria Broth with kanamycin (100 μg/mL) was inoculated with *E. coli* UT5600/pBBTH-HHHR, with *E. coli* UT5600/pBBTH as a control, at 37^{o}C, 250 rpm until an OD600 between 0.3 - 0.4. The culture was then induced with 0.5 mM IPTG and incubated for 20 hours at 22^{o}C, 200 rpm. Cells were harvested using centrifugation (3500 x *g*, 40 minutes, 4^{o}C), washed once with Tris-buffered Saline (TBS), and re-suspended in TBS with 1 mM $CaCl_2$, concentrated ten times the original volume. To construct xylanosomes on the cell surface, 1 mL of concentrated cells were incubated with appropriate combinations of purified X10Doc, X11Doc, FAEDoc, and/or ABFDoc in 20 mL scintillation vials at 25^{o}C for 1 hour at 125 rpm. Cells were then transferred to 1.5-mL microcentrifuge tubes and collected with centrifugation (5000 x *g*, 2 minutes). Supernatant was removed and cells re-suspended in 200 μL TBS buffer with 1 mM $CaCl_2$. Optical density was measured at 600 nm to characterize enzymatic activities.

Xylanase activity of constructed cell-surface xylanosomes on oat spelt xylan was performed by combining 100 μL cells with 100 μL 1% oat spelt xylan in 10 mM Tris-

HCl buffer, pH 7.0, and incubated at 50^{o}C for 30 minutes at 1000 rpm using a thermomixer. Cells were harvested (16000 x *g*, 2 minutes) after boiling the reactions for 5 minutes, and the supernatant collected for released reducing sugar, as mentioned previously in Section 3.5.4. Ferulic acid esterase activity was measured by combining 200 mL cells with 1 mM methyl ferulate (Apin Chemicals, UK) in 10 mM Tris-HCl buffer, pH 7.0 and incubating the mixture at 40^{o}C for 30 minutes, 1000 rpm using a thermomixer. The reaction was stopped by adding glacial acetic acid and the supernatant collected using centrifugation (16000 x *g*, 2 minutes). Ferulic acid concentrations were measured as mentioned previously in Section 3.5.4.

3.6 References

1. Saha, B.C., et al., *Lignocellulose biodegradation*. ACS symposium series 889. 2004, Washington, DC: American Chemical Society : Distributed by Oxford University Press. xii, 400 p.

2. Cherry, J.R. and A.L. Fidantsef, *Directed evolution of industrial enzymes: an update.* Current Opinion in Biotechnology, 2003. **14**(4): p. 438-443.

3. Sorensen, H.R., et al., *Enzymatic hydrolysis of wheat arabinoxylan by a recombinant "minimal" enzyme cocktail containing beta-xylosidase and novel endo-1,4-beta-xylanase and alpha-l-arabinofuranosidase activities.* Biotechnol Prog, 2007. **23**(1): p. 100-7.

4. Fontes, C.M. and H.J. Gilbert, *Cellulosomes: highly efficient nanomachines designed to deconstruct plant cell wall complex carbohydrates.* Annu Rev Biochem, 2010. **79**: p. 655-81.

5. Adams, J.J., et al., *Mechanism of bacterial cell-surface attachment revealed by the structure of cellulosomal type II cohesin-dockerin complex.* Proc Natl Acad Sci U S A, 2006. **103**(2): p. 305-10.

6. Haimovitz, R., et al., *Cohesin-dockerin microarray: Diverse specificities between two complementary families of interacting protein modules.* Proteomics, 2008. **8**(5): p. 968-79.

7. Jindou, S., et al., *Cohesin-dockerin interactions within and between Clostridium josui and Clostridium thermocellum: binding selectivity between cognate dockerin and cohesin domains and species specificity.* J Biol Chem, 2004. **279**(11): p. 9867-74.

8. Mechaly, A., et al., *Cohesin-dockerin interaction in cellulosome assembly: a single hydroxyl group of a dockerin domain distinguishes between nonrecognition and high affinity recognition.* J Biol Chem, 2001. **276**(13): p. 9883-8.

9. Pages, S., et al., *Species-specificity of the cohesin-dockerin interaction between Clostridium thermocellum and Clostridium cellulolyticum: prediction of specificity determinants of the dockerin domain.* Proteins, 1997. **29**(4): p. 517-27.

10. Fierobe, H.P., et al., *Design and production of active cellulosome chimeras. Selective incorporation of dockerin-containing enzymes into defined functional complexes.* J Biol Chem, 2001. **276**(24): p. 21257-61.

11. Fierobe, H.P., et al., *Degradation of cellulose substrates by cellulosome chimeras. Substrate targeting versus proximity of enzyme components.* J Biol Chem, 2002. **277**(51): p. 49621-30.

12. Fierobe, H.P., et al., *Action of designer cellulosomes on homogeneous versus complex substrates: controlled incorporation of three distinct enzymes into a defined trifunctional scaffoldin.* J Biol Chem, 2005. **280**(16): p. 16325-34.

13. Mingardon, F., et al., *Incorporation of fungal cellulases in bacterial minicellulosomes yields viable, synergistically acting cellulolytic complexes.* Appl Environ Microbiol, 2007. **73**(12): p. 3822-32.

14. Vazana, Y., et al., *Interplay between Clostridium thermocellum family 48 and family 9 cellulases in cellulosomal versus noncellulosomal states.* Appl Environ Microbiol. **76**(10): p. 3236-43.

15. Huang, Y.H., C.T. Huang, and R.S. Hseu, *Effects of dockerin domains on Neocallimastix frontalis xylanases.* FEMS Microbiol Lett, 2005. **243**(2): p. 455-60.

16. Fierobe, H.P., et al., *Cellulosome from Clostridium cellulolyticum: molecular study of the Dockerin/Cohesin interaction.* Biochemistry, 1999. **38**(39): p. 12822-32.

17. Pinheiro, B.A., et al., *The Clostridium cellulolyticum dockerin displays a dual binding mode for its cohesin partner.* J Biol Chem, 2008. **283**(26): p. 18422-30.

18. Schaeffer, F., et al., *Duplicated dockerin subdomains of Clostridium thermocellum endoglucanase CelD bind to a cohesin domain of the scaffolding protein CipA with distinct thermodynamic parameters and a negative cooperativity.* Biochemistry, 2002. **41**(7): p. 2106-14.

19. Agrawal, M., Z. Mao, and R.R. Chen, *Adaptation yields a highly efficient xylose-fermenting Zymomonas mobilis strain.* Biotechnol Bioeng.

20. Biely, P., et al., *Endo-beta-1,4-xylanase families: differences in catalytic properties.* J Biotechnol, 1997. **57**(1-3): p. 151-66.

21. Shin, H.D., et al., *A complete enzymatic recovery of ferulic acid from corn residues with extracellular enzymes from Neosartorya spinosa NRRL185.* Biotechnol Bioeng, 2006. **95**(6): p. 1108-15.

22. Blum, D.L., et al., *Feruloyl esterase activity of the Clostridium thermocellum cellulosome can be attributed to previously unknown domains of XynY and XynZ.* J Bacteriol, 2000. **182**(5): p. 1346-51.

23. Gold, N.D. and V.J. Martin, *Global view of the Clostridium thermocellum cellulosome revealed by quantitative proteomic analysis.* J Bacteriol, 2007. **189**(19): p. 6787-95.

24. Raman, B., et al., *Impact of pretreated Switchgrass and biomass carbohydrates on Clostridium thermocellum ATCC 27405 cellulosome composition: a quantitative proteomic analysis.* PLoS One, 2009. **4**(4): p. e5271.

25. Tsai, S.L., et al., *Functional assembly of minicellulosomes on the Saccharomyces cerevisiae cell surface for cellulose hydrolysis and ethanol production.* Appl Environ Microbiol, 2009. **75**(19): p. 6087-93.

26. Wen, F., J. Sun, and H. Zhao, *Yeast surface display of trifunctional minicellulosomes for simultaneous saccharification and fermentation of cellulose to ethanol.* Appl Environ Microbiol. **76**(4): p. 1251-60.

27. Wang, B., B. Cheng, and H. Feng, *Enriched arabinoxylan in corn fiber for value-added products.* Biotechnol Lett, 2008. **30**(2): p. 275-9.

28. Miller, G.L., *Use of dinitrosaliclic acid reagen for determination of reducing sugars.* Analytical Chemistry, 1959(31): p. 127-132.

29. Shin, H.D. and R.R.Z. Chen, *Production and characterization of a type B feruloyl esterase from Fusarium proliferatum NRRL 26517.* Enzyme and Microbial Technology, 2006. **38**(3-4): p. 478-485.

CHAPTER 4

CLONING AND CHARACTERIZATION OF *FEE1B* FROM *CELLVIBRIO JAPONICUS*[1]

4.1 Abstract

Ferulic acid esterases (FAE) are important enzymes with potential applications in enzymatic release of ferulic acid and lignocellulose hydrolysis. Fee1B is a putative FAE identified from a recently sequenced genome of *Cellvibrio japonicas*. The multi-domain enzyme is highly unusual in its presence of two tandem N-terminal carbohydrate binding modules (CBMs) belonging to family 2 and 35. The purified Fee1B has a pH and temperature optima of 6.5 and 35-40^{o}C, respectively. Fee1B can be classified as a type-D FAE with broad substrate specificity, active on multiple methyl esters of hydroxycinnamic acid and *p*-nitrophenyl acetate. The two CBMs were apparently crucial for enzyme activities as the truncated enzyme significantly reduced the ferulic acid esterase activity. The potential of Fee1B in biomass hydrolysis was demonstrated with a complex xylan substrate, wheat arabinoxylan. Fee1B enhanced xylanase-catalyzed release of reducing sugars by as much as 8 times as xylanase alone. Furthermore, Fee1B alone was able to release ferulic acid from several biomass substrates, most significantly

---

[1] Portions of this chapter were previously published : McClendon, S., Shin, H-D., Chen, R. (2011) "A novel bacterial ferulic acid esterase from *Cellvibrio japonius* and its application in ferulic acid release and xylan hydrolysis." Biotechnology Letters 33(1): 47-54.

from destarched corn bran, liberating 4.7 nmol ferulic acid/mg insoluble substrate. Synergy with xylanase in ferulic acid release, however, varies with substrates, with high synergy observed with corn bran, moderate with destarched corn bran, and negative synergy with wheat arabinoxylan. The data presented here provide the first experimental evidence for ferulic acid esterase activity of Fee1B. Its broad substrate specificity and activity toward complex xylan substrates should allow the enzyme to find applications in ferulic acid release as well as in biomass utilization.

4.2 Introduction

Ferulic acid occurs naturally within the cell wall of numerous types of plants and crops [1] and has numerous applications in the food, cosmetic, and pharmaceutical industries due to its antioxidant properties [2]. Crop residues provide a natural source of ferulic acid for use in various chemical or biological processes, such as the bioproduction of vanillin [3]. The enzyme responsible for the release of ferulic acid from biomass is feruloyl esterase (E.C. 3.1.1.73, FAE), which catalyzes the hydrolysis of the ester bond between hydroxycinnamic acids and sugars [4]. FAEs could be used synergistically with other hemicellulases for release of ferulic acid. Indeed, this was observed in several studies [5, 6].

FAE, as a hemicellulase, is also useful for lignocellulose hydrolysis. Xylan, for example, is a β-1,4-linked xylose polymer with various side chains of arabinose, acetic acid, glucuronic acid, and phenolic acids, such as ferulic acid, which are ester-linked to arabinose side groups. Cross-linking ferulic acids between xylan chains or to lignin significantly increases biomass recalcitrance [7, 8]. Thus, removal of ferulic acid from plant cell walls is important in effectively hydrolyzing biomass by decreasing substrate

complexity and increasing cellulose accessibility. In nature, the *Clostridium thermocellum* cellulosome system contains FAE activity in the multi-protein structure, suggesting an important role of FAE in biomass degradation [9].

Due to these important applications, discovery and characterization of novel FAEs continue to be an important area of research [2, 4]. Most well-characterized FAEs are fungal enzymes [10] which usually require eukaryotic expression systems for recombinant expression. Bacterial FAEs can be made accessible through recombinant expression using well established *Escherichia coli* systems and are suitable for metabolic engineering of bacterial strains for consolidated bioprocessing. A search of the recently sequenced genome of cellulolytic *Cellvibrio japonicus* (formerly *Pseudomonas fluorescens* subp. *cellulosa*) identifies two FAE genes, *fee1A* and *fee1B* (notation according to NCBI). The two genes are highly unusual in the presence of two different carbohydrate binding modules (CBMs) in their N-terminus. While Fee1A was previously reported [11], Fee1B was not known to exist prior to genome sequencing and has not been described. In this work, we report the cloning and expression of the putative feruloyl esterase gene, *fee1B,* using an *E. coli* expression system. The corresponding protein is purified to homogeneity and characterized with respect to substrate specificity, and pH and temperature profiles. Additionally, its synergistic use with other hemicellulases in both ferulic acid release and biomass hydrolysis is investigated with multiple complex biomass substrates.

4.3 Results

4.3.1 *fee1B* encodes a ferulic acid esterase with two tandem CBMs at the N-terminus

Cellvibrio japonicus is a cellulolytic bacterium capable of degrading all major plant cell wall polysaccharides, including crystalline cellulose, mannan, and xylan [12]. Its genome encodes 130 glycosyl hydrolases, including several hemicellulases. In particular, the genome sequencing uncovers the presence of two ferulic acid esterases of high homology to each other. Although one of the FAEs, XylD (encoded by *fee1* according to NCBI, or *fae1* according to Deboy in 2008), was reported and extensively studied [5, 11], the second gene *fee1B*, which resides upstream from *fee1A* has not been described before. Besides the catalytic domain, *fee1B* has two CBMs in tandem at the N-terminus, which are classified as family 2 and family 35, respectively. Three other genes in the *C. japonicus* genome, a xylanase (*xyn10B*), arabinofuranosidase (*abf62A*), and *fee1A*, also have the same CBM2-CBM35 domains at the N-terminus of their respective sequence, suggesting their prominent role in lignocellulose degradation. These genes are clustered together in the genome and *fee1B* is found four ORFs upstream from *fee1A*. These hemicellulolytic genes appear to be regulated by L-arabinose or perhaps another 5-carbon sugar via putative AraC transcription regulators [13] named CJA_3278 (*xyn10B*, *abf62A*, & *fee1B*) and CJA_3284 (*ebg98* & *fee1A*). The genome location and modular composition of *fee1B* is shown in Figure 4.1.

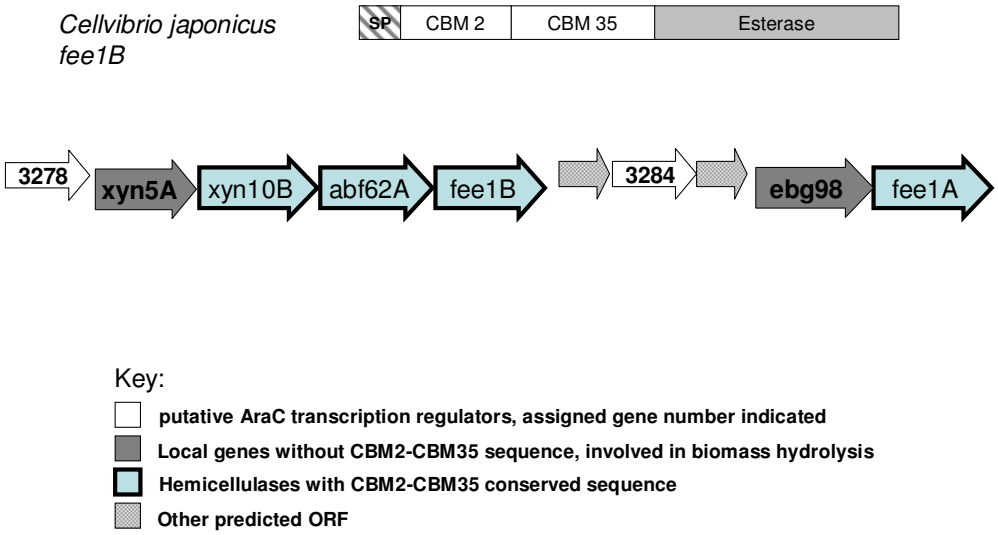


Figure 4.1 Domain structure of Fee1B and its location within a hemicellulase gene cluster in genome of *C. japonicus*

A comparison of the two FAE genes showed 88% identity and 92% homology over the entire amino acid (AA) sequence. The putative esterase domains were shown to have 86% homology with each other, each containing a G-X-S-X-G motif typical of serine esterases. Both genes contain an identical sequence of 299 amino acids following the predicted signal peptides (SP), which includes the tandem CBM2-CBM35 AA sequence. Out of all sequenced and characterized FAEs (carbohydrate esterase family 1), Fee1A and Fee1B are the only reported bacterial ferulic acid esterases with two tandem CBMs from different families.

Table 4.1 Purification summary of recombinant Fee1B protein

Step	Total Activity (mU)	Total Protein (mg)	Specific Activity (mU/mg)	Purification Fold	Recovery (%)
Soluble Extract	23.8	10.76	2.2	1	100
IMAC	14.6	0.55	26.6	12	61
Dialysis	14.3	0.53	27.0	12	60

To confirm the FAE activity of fee1B, the gene was amplified using PCR and cloned into *E. coli* JM109 via pET20b(+) as described in Materials and Methods. Fee1B was successfully expressed in *E. coli* BL21(DE3) as a His-tagged fusion protein and purified using immobilized metal affinity chromatography. The one-step affinity chromatography yielded a purified protein with an expected molecular weight of 61 kDa and an overall recovery of 60%, as listed in Table 4.1. The purified recombinant enzyme was active toward a common FAE synthetic substrate, methyl ferulate, with about 28 mU/mg protein, indicating that the gene *fee1B* encodes a functional FAE.

4.3.2 Enzyme characterization

To further characterize the enzyme, other synthetic substrates were used. Table 4.2 shows that in addition to methyl ferulate, the enzyme was also active toward other hydroxycinnamic substrates tested. The activity order was as follows: methyl ferulate > methyl *p*-coumarate > methyl sinapate > methyl caffeate, the same substrate preference reported for fee1A. Based on this profile, the enzyme can be classified as type-D FAE [14]. Besides typical FAE activity, the enzyme exhibited high activity toward *p*-

nitrophenyl acetate, showing it has acetylxylan esterase activity as well. This was unusual for a typical FAE but was observed with a previously characterized fee1A [11].

Table 4.2 Substrate specificity of Fee1B from *C. japonicus*

Substrate	**Specific Activity[a] (mU/mg protein)**	
	Fee1B	**Truncated Fee1B**
p-Nitrophenyl acetate	50 ± 2	36 ± 2
Methyl ferulate	27.5 ± 1.7	4.8 ± 1.3
Methyl *p*-coumarate	12.4 ± 0.8	2.9 ± 0.9
Methyl caffeate	5.4 ± 0.3	1.1 ± 1.0
Methyl sinapate	10.5 ± 1.9	3.8 ± 1.1

[a] Each reaction was performed in duplicate.

To assess the role of two tandem CBMs, the truncated enzyme was obtained by cloning only the C-terminal catalytic domain into the *E. coli* expression system. As shown in Table 4.2, removal of CBMs caused a significant decrease in FAE activities, where up to 80% activity was lost as a result. Reasons for this loss of activity are unclear, but may relate to changes in protein folding or substrate binding due to lack of CBM modules. Interestingly, the reduction of esterase activity on *p*-nitrophenyl acetate was not as substantial, decreasing activity by about 30%. This observation is in sharp contrast with what was observed with FeelA, where removal of the N-terminal CBM completely abolished the activity [11].

Although the esterase domain of fee1B showed 71% AA sequence homology with putative glycosyl hydrolase family 62 proteins from *Micromonospora* sp. (annotated as alpha-arabinofuranosidases), Fee1B was not active against *p*-nitrophenyl alpha-L-arabinofuranoside.

The enzyme was further characterized with respect to its pH and temperature profiles. As shown in Figure 4.2 and Figure 4.3, the optimal pH and temperature of Fee1B is 6.5 and between 35-40^{o}C, respectively. This is largely expected from a mesophilic bacterial enzyme. Fee1B is fairly stable over a wide range of pH and is stable at temperature up to 40^{o}C. The enzyme lost about 30% activity upon incubation at 45^{o}C for 30 minutes.

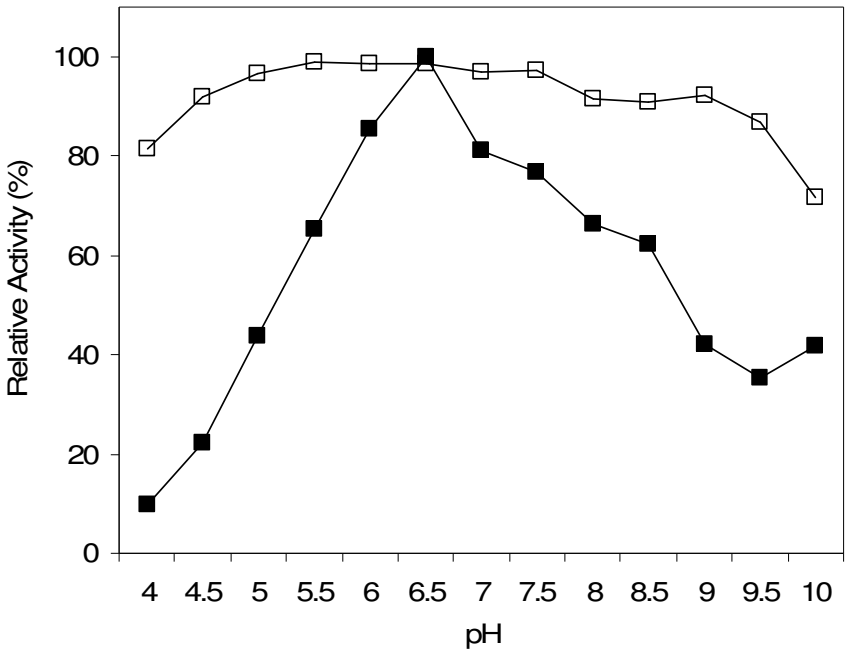


Figure 4.2 Effect of pH on activity (closed square) and stability (open square) of Fee1B from *C. japonicus*. Activity measured by incubating enzyme in appropriate buffer with methyl ferulate as substrate. Stability was determined by storing the enzyme at various pH conditions for 16 hr at 4^{o}C and measuring residual activity. Relative activity in percentage (%). Absolute activity = 41 mU/mg protein.

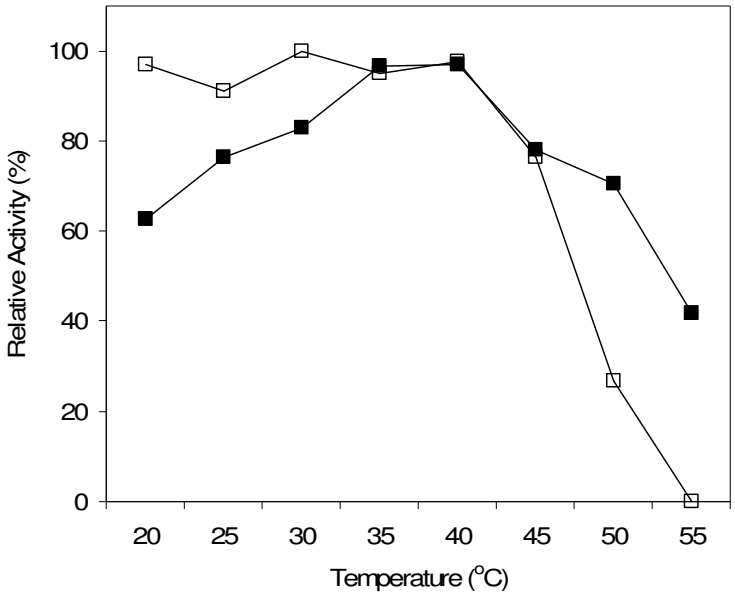


Figure 4.3 Effect of temperature on activity (closed square) and stability (open square) of Fee1B from *C. japonicus*. Activity measured by incubating enzyme at designated temperature with methyl ferulate as substrate at pH 6. Stability was determined by incubating the enzyme at various temperatures and measuring residual activity. Relative activity in percentage (%). Absolute activity = 41 mU/mg protein.

4.3.3 Use of Fee1B in ferulic acid release and hemicellulose hydrolysis

As removal of ferulic acid crosslinking could potentially reduce the recalcitrance of hemicellulose, Fee1B may be used synergistically with other hemicellulases to enhance hydrolysis of hemicelluloses substrate. This was tested using wheat arabinoxylan (WAX) as substrate. When Fee1B was combined with either a glycosyl hydrolyase (GH) family 10 or family 11 xylanase (Xyl10, Xyl11), synergistic effects were observed for the release of soluble sugars from wheat arabinoxylan (WAX). Compared to the single xylanase reactions, both Xyl10 + Fee1B and Xyl11 + Fee 1B enzyme combinations

showed a 5- and 8-fold increase, respectively, in the total amount of released reducing sugars from WAX, with the highest concentration being 4.9 mM or 0.49 μmol sugars/mg insoluble substrate (Figure 4.4). Synergy of Fee1B with xylanase family 10 or xylanase family 11 are 2.2 and 1.8, respectively. These results indicate the removal of ester-linked ferulate side chains improves hydrolysis of hemicellulose.

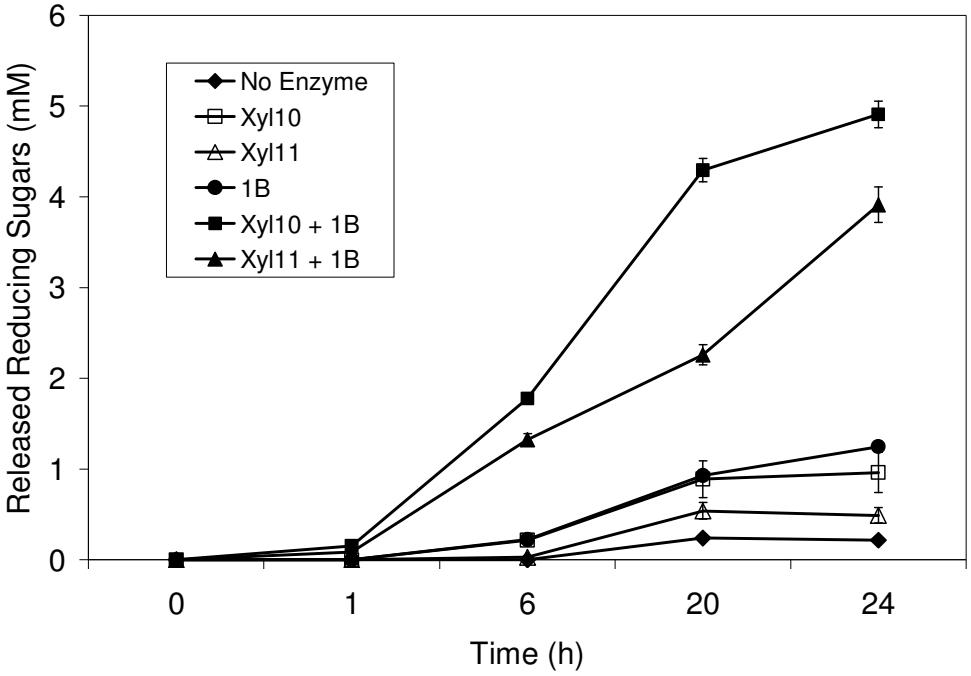


Figure 4.4 Release of reducing sugars from wheat arabinoxylan over time. Data shown were averages of three independent experiments. Error bars represent standard deviation.

Use of Fee1B for enzymatic recovery of ferulic acid was also tested. Fee1B alone can release significant amount of ferulic acid (FA) from WAX (Figure 4.5 and Table 4.3). At a loading of 20 mU/mg substrate, Fee1B can cleave 0.41 nmol ferulic acid/mg insoluble WAX or 2.8% recovery of FA. In contrast to the release of reducing sugar, however, the addition of either Xyl10 or Xyl11 did not improve the release of FA from

WAX even for extended reactions. Figure 5 shows a slight improvement in ferulic acid release at early time points with addition of either xylanase but the advantage disappears after six hours.

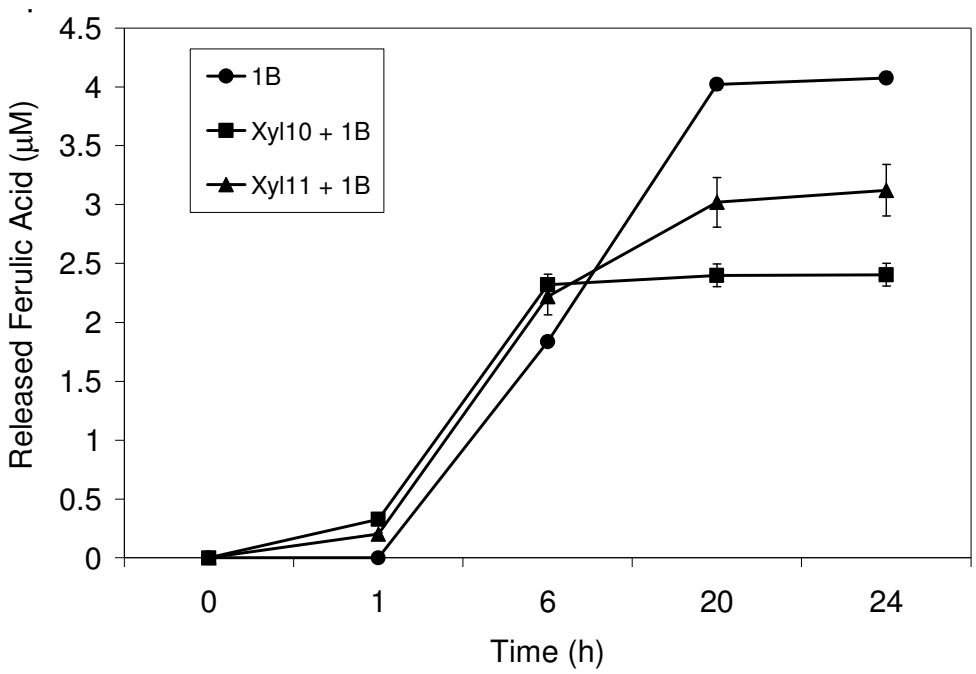


Figure 4.5 Release of ferulic acid from wheat arabinoxylan over time. Data shown were averages of three independent experiments. Error bars represent standard deviation.

Table 4.3 Ferulic acid release from various complex xylan substrates using Fee1B and xylanase [b]

Enzyme	Wheat Arabinoxylan		Corn Bran		Destarched Corn Bran	
	Released Ferulic Acid (nmol/mg substrate)	Synergy[c]	Released Ferulic Acid (nmol/mg substrate)	Synergy	Released Ferulic Acid (nmol/mg substrate)	Synergy
Fee1B	0.41 ± 0.01	n/a[d]	0.48 ± 0.09	n/a	4.74 ± 0.72	n/a
Xyl10 + Fee1B	0.24 ± 0.01	0.59	0.60 ± 0.01	1.25	5.19 ± 0.78	1.09
Xyl11 + Fee1B	0.31 ± 0.02	0.43	0.94 ± 0.14	1.93	5.21 ± 0.79	1.10

[b] Each experiment performed in triplicate.
[c] Synergy is calculated dividing the product released of enzyme mixtures by the sum of product released by each individual enzyme.
[d] n/a – not applicable

The use of Fee1B to release ferulic acid was also tested on corn bran and destarched corn bran, substrates with higher ferulic acid content than WAX. Apparently, destarching corn bran greatly improved the enzymatic release of ferulic acid, increasing ferulic acid released from destarched corn bran 10 times compared to corn bran without destarching, as listed in Table 4.3. Released ferulic acid concentrations increased from 0.48 nmol/mg substrate to 0.6 nmol/mg substrate and 0.94 nmol/mg substrate with Xyl10 and Xyl11, respectively, on corn bran. The effect of xylanase addition varied with both the substrates and type of xylanase used. Significant synergy (up to 1.9) was observed with corn bran whereas only slight synergy (synergy factor of 1.1) was observed with destarched corn bran. The highest synergy was observed with xylanase type 11 and corn

bran (Table 4.3) whereas the highest release of FA was observed on destarched corn bran at 5.21 nmol/mg substrate, using Xyl11 and Fee1B. Taken together, these results suggest that Fee1B could be used for ferulic acid release from various substrates but the synergy with xylanase and possibly other hemicellulases varies greatly with substrates used.

4.4 Discussion

A putative ferulic acid esterase gene, *fee1B*, from *Cellvibrio japonicus* was cloned and expressed in *E. coli* and purified for enzyme characterization and application. The enzyme is unusual in the presence of two CBMs belonging to family 2 and family 35, respectively. Family 2 CBM has known affinity for cellulose, while family 35 CBM binds to amorphous cellulose or xylan. The roles of the CBMs in Fee1B appear to greatly contribute to the overall activity of the enzyme. When both CBMs were removed, the activity of Fee1B decreased 2 to 5-fold on methyl ferulate, methyl *p*-coumarate, and methyl caffeate, and methyl sinapate. Numerous esterases (both ferulic acid and acetylxylan esterases) have CBMs as a part of their structure. Bacterial FAEs that contain CBMs typically possess a single CBM. Fee1B and Fee1A are the only FAEs that have two CBMs. Although family 35 CBM is closely related to family 6 CBMs seen in other bacterial FAEs, the family 2 CBM of Fee1B is not typically found in hemicellulases. Since *C. japonicus* is not predicted to produce cellulosomes, multiple CBMs could be needed to place its enzymes in close proximity to the target substrate for efficient hydrolysis.

Fee1B displays activity on multiple synthetic and complex substrates. It has the highest activity on *p*-nitrophenyl acetate, followed by methyl ferulate as the preferred

hydroxycinnamic acid substrate, suggesting dual-functionality in releasing both acetyl and ferulate groups. Fee1B also shows the ability to cleave ferulic acid from wheat arabinoxylan, corn bran and destarched corn bran, suggesting potential application in hydrolysis of a broad range of hemicellulose substrate. By comparing its amino acid sequence to known fungal and bacterial FAEs and its broad substrate specificity, Fee1B is classified as a type-D ferulic acid esterase, which includes fungal esterases, such as EstA from *Piromyces equi* [15] and NcFaeD from *Neurospora crassa* [16], and Fee1A from the same organism [14].

The data presented here clearly shows that Fee1B could be synergistically used with other hemicellulases to improve the hydrolysis of plant biomass. Released reducing sugar (RRS) concentration increased significantly by combining Fee1B with either Xyl10 or Xyl11. Although the highest concentration of RRS was observed with the Xyl10+Fee1B combination, the greatest synergy was seen with the Xyl11+Fee1B combination. Since family 11 xylanases prefer less substituted regions of the xylan backbone [17], the removal of ferulic acid and/or acetyl groups exposes more cleavage sites for Xyl11, therefore enhancing its activity. Family 10 xylanases are more active on more substituted regions of xylan, explaining its higher activity on WAX overall [17]. In addition to its synergy with other hemicellulases, the pH and temperature profile for Fee1B suggest that it could be applied in simultaneous saccharification and fermentation (SSF) processes or consolidated bioprocessing, where mesophilic microorganisms are used.

In conclusion, we confirm that fee1B gene encodes a second functional ferulic acid esterase in *Cellvibrio japonicus*. Fee1B is an unusual bacterial FAE with two

different CBMs in tandem, which appear to be crucial to its activity. Fee1B is a type-D FAE with broad substrate specificity and can be synergistically used with other hemicellulases in both ferulic acid release and lignocellulose hydrolysis. With high compatibility with other mesophilic bacterial enzymes, it will be useful for consolidated bioprocessing or SSF.

4.5 Materials and Methods

4.5.1 Strains and materials

The genomic DNA of *Cellvibrio japonicus* NCIMB 10462 was isolated using a fungal/bacterial genomic DNA extraction kit from ZymoResearch (Orange, CA, USA) after growth of the strain in Luria broth (LB). *Escherichia coli* JM109 and BL21(DE3) strains were used in cloning and expression experiments, respectively. Wheat arabinoxylan (Megazyme, Ireland), corn bran (SunOpta, Bedford, MA, USA), and destarched corn bran were complex substrates used for release of ferulic acid and enzymatic hydrolysis. Destarched corn bran was prepared according to Wang [18] using Termayl (Novozymes, Denmark) instead of G-ZYME.

4.5.2 Cloning and expression

Primers Fee1B-F (5'– ATCT*CATATG*CATCGGGTTAATTGGAG – 3', *Nde*I site) and Fee1B-R (5' – TCTAAT*CTCGAG*GAACTGTGTGAAGAATTGCCAG – 3', *Xho*I site) were used to amplify the full *fee1B* gene from the genome of *C. japonicus* via PCR. The gene fragment was ligated into *Nde*I-*Xho*I-linearized pET20B(+) to obtain the pEfee1B, which was used to transform *E. coli* JM109 for plasmid propagation. *E. coli* BL21(DE3) was transformed with pEfee1B for expression of Fee1B. The truncated

version of Fee1B (without the two CBMs) was cloned in a similar manner as described for the full-length gene except the following primers were used for amplification the corresponding gene, Fee1Bfae-F (5'–ATTGAA*GGATCC*GGTGCCGCAGTACCCACTG – 3', *Bam*HI site) and Fee1Bfae-R (5'-AGTAGG*AAGCTT*GAATTGTGTGAAGAATTGCCAGG-3', *Hind*III site). The resultant recombinant strains were cultured in LB media with 100 μg/mL ampicillin at 30^{o}C and induced with 0.2 mM IPTG at 16^{o}C for 24 hours for Fee1B or truncated Fee1B expression. After re-suspending the cell pellet in 50 mM sodium phosphate, 0.3 M NaCl buffer, pH 8.0, the cells were sonicated and the soluble fraction collected using centrifugation. The target protein was purified using immobilized metal (nickel) affinity chromatography (HIS-Select™ Resin, Sigma-Aldrich) by eluting the protein from the resin with 50 mM sodium phosphate, 0.3 M NaCl, 250 mM imidazole buffer, pH 8.0, and dialyzed against 10 mM Tris-HCl, pH 7.5 for removal of imidazole for subsequent assays.

4.5.3 Enzyme assays

Ferulic acid esterase (FAE) activity was determined by measuring the amount of ferulic acid released from methyl ferulate (Apin Chemicals, Oxon, UK). Reactions were carried out by adding enzyme to 100 mM sodium phosphate buffer, pH 7.0, with 1 mM methyl ferulate and incubating the mixture at 40^{o}C for 30 minutes. The reaction was stopped by adding glacial acetic acid. FAE activity on complex substrates was performed by adding 20 mU/mg of enzyme to 10 mg/mL wheat arabinoxylan (WAX) or (destarched) corn bran (CB, DCB) and incubating at 40^{o}C with and without presence of xylanase (0.1 U/mg substrate) in a thermomixer for 24 hours at 1400 rpm. The reaction

was ended by boiling the mixture for 5 minutes and collecting the supernatant via centrifugation. Family 10 and 11 xylanases are from *Clostridium thermocellum* [19] and *Bacillus halodurans* [20], respectively. Both xylanases were cloned into and recombinantly expressed in *E. coli* and purified via histidine tags. One unit of activity is defined as the amount of enzyme that releases 1 μmol of ferulic acid per min of reaction. Activity on other methyl esters of hydroxycinnamic acids was performed the same as FAE activity, except the substrate was varied to methyl sinapate, methyl caffeate, or methyl *p*-coumarate (Apin Chemicals, Oxon, UK).

Esterase activity was assayed by incubating the enzyme in 100 mM sodium phosphate buffer, pH 6.0 with 1 mM *p*-nitrophenyl acetate at 37°C for 10 minutes, and adding 0.5 M NaOH to end the reaction. Quantification of activity was done by measuring the release of *p*-nitrophenol via absorbance at 410 nm.

4.5.4 Enzyme characterization

To determine an optimal pH for Fee1B, FAE activity was measured as described above except the reaction was carried out in a pH range of 4.0 – 10.0 using the following buffers: 50 mM acetate-phoshpate buffer (pH 4.5 – 7.0), 50 mM Tris-HCl buffer (pH 7.5 – 8.5), and 50 mM glycine-NaOH buffer (pH 9.0 – 10.0). pH stability was determined by measuring the residual FAE activity (pH 7.0, 40°C) after incubating the enzyme in an appropriate buffer for 16 hours at 4°C. Optimal temperature was determined by performing the FAE activity assay at various temperatures, ranging from 20 - 50°C. Thermostability was determined by incubating the enzyme at the designated temperature (20 – 50°C) for 30 minutes and proceeding to perform the FAE activity assay in 100 mM sodium phosphate buffer, pH 7.0, at 40°C.

4.5.5 Analytical methods

Protein concentration was measured by the method described by Bradford using a protein reagent dye (Bio-Rad), with bovine serine albumin as the standard. Reducing sugars released from complex substrates were measured using the method described by Miller [21] with xylose as the standard. SDS-PAGE was performed using a 12% Tris-HCl gel (Bio-Rad) and stained with Bio-safe Coomassie Blue (Bio-Rad). Quantification of released ferulic acid was performed by high performance liquid chromatography as described by Shin and Chen [22].

4.5 References

1. Mathew, S. and T.E. Abraham, *Ferulic acid: an antioxidant found naturally in plant cell walls and feruloyl esterases involved in its release and their applications.* Crit Rev Biotechnol, 2004. **24**(2-3): p. 59-83.

2. Fazary, A.E. and Y.H. Ju, *Feruloyl esterases as biotechnological tools: current and future perspectives.* Acta Biochim Biophys Sin (Shanghai), 2007. **39**(11): p. 811-28.

3. Shin, H.D., et al., *A complete enzymatic recovery of ferulic acid from corn residues with extracellular enzymes from Neosartorya spinosa NRRL185.* Biotechnol Bioeng, 2006. **95**(6): p. 1108-15.

4. Wong, D.W., *Feruloyl esterase: a key enzyme in biomass degradation.* Appl Biochem Biotechnol, 2006. **133**(2): p. 87-112.

5. Bartolome, B., et al., *An Aspergillus niger esterase (ferulic acid esterase III) and a recombinant Pseudomonas fluorescens subsp cellulosa esterase (XylD) release a 5-5' ferulic dehydrodimer (diferulic acid) from barley and wheat cell walls.* Applied and Environmental Microbiology, 1997. **63**(1): p. 208-212.

6. Faulds, C.B., et al., *Synergy between xylanases from glycoside hydrolase family 10 and family 11 and a feruloyl esterase in the release of phenolic acids from cereal arabinoxylan.* Appl Microbiol Biotechnol, 2006. **71**(5): p. 622-9.

7. Grabber, J.H., R.D. Hatfield, and J. Ralph, *Diferulate cross-links impede the enzymatic degradation of non-lignified maize walls.* Journal of the Science of Food and Agriculture, 1998. **77**(2): p. 193-200.

8. Grabber, J.H., J. Ralph, and R.D. Hatfield, *Ferulate cross-links limit the enzymatic degradation of synthetically lignified primary walls of maize.* Journal of Agricultural and Food Chemistry, 1998. **46**(7): p. 2609-2614.

9. Blum, D.L., et al., *Feruloyl esterase activity of the Clostridium thermocellum cellulosome can be attributed to previously unknown domains of XynY and XynZ.* J Bacteriol, 2000. **182**(5): p. 1346-51.

10. Koseki, T., et al., *Occurrence, properties, and applications of feruloyl esterases.* Applied Microbiology and Biotechnology, 2009. **84**(5): p. 803-810.

11. Ferreira, L.M., et al., *A modular esterase from Pseudomonas fluorescens subsp. cellulosa contains a non-catalytic cellulose-binding domain.* Biochem J, 1993. **294 (Pt 2)**: p. 349-55.

12. DeBoy, R.T., et al., *Insights into plant cell wall degradation from the genome sequence of the soil bacterium Cellvibrio japonicus.* J Bacteriol, 2008. **190**(15): p. 5455-63.

13. Gallegos, M.T., et al., *Arac/XylS family of transcriptional regulators.* Microbiol Mol Biol Rev, 1997. **61**(4): p. 393-410.

14. Crepin, V.F., C.B. Faulds, and I.F. Connerton, *Functional classification of the microbial feruloyl esterases.* Appl Microbiol Biotechnol, 2004. **63**(6): p. 647-52.

15. Fillingham, I.J., et al., *A modular cinnamoyl ester hydrolase from the anaerobic fungus Piromyces equi acts synergistically with xylanase and is part of a multiprotein cellulose-binding cellulase-hemicellulase complex.* Biochem J, 1999. **343 Pt 1**: p. 215-24.

16. Crepin, V.F., C.B. Faulds, and I.F. Connerton, *Identification of a type-D feruloyl esterase from Neurospora crassa.* Appl Microbiol Biotechnol, 2004. **63**(5): p. 567-70.

17. Biely, P., et al., *Endo-beta-1,4-xylanase families: differences in catalytic properties.* J Biotechnol, 1997. **57**(1-3): p. 151-66.

18. Wang, B., B. Cheng, and H. Feng, *Enriched arabinoxylan in corn fiber for value-added products.* Biotechnol Lett, 2008. **30**(2): p. 275-9.

19. Hayashi, H., et al., *Sequence of xynC and properties of XynC, a major component of the Clostridium thermocellum cellulosome.* J Bacteriol, 1997. **179**(13): p. 4246-53.

20. Wamalwa, B.M., et al., *High-level heterologous expression of Bacillus halodurans putative xylanase xyn11a (BH0899) in Kluyveromyces lactis.* Biosci Biotechnol Biochem, 2007. **71**(3): p. 688-93.

21. Miller, G.L., *Use of dinitrosaliclic acid reagen for determination of reducing sugars.* Analytical Chemistry, 1959(31): p. 127-132.

22. Shin, H.D. and R.R.Z. Chen, *Production and characterization of a type B feruloyl esterase from Fusarium proliferatum NRRL 26517.* Enzyme and Microbial Technology, 2006. **38**(3-4): p. 478-485.

CHAPTER 5

DESIGN, CONSTRUCTION, AND APPLICATION OF THREE UNIT XYLANOSOMES

5.1 Abstract

Complete hydrolysis of xylan requires six different enzymatic activities. Previously constructed xylanosomes only allowed one xylanase and one other accessory enzyme for xylan hydrolysis, combining up to three different enzymatic activities. This chapter discusses the molecular design, construction, and application of three-unit xylanosomes, which include five out of the six required enzymes: xylanase, arabinofuranosidase, xylosidase, ferulic acid esterase, and acetylxylan esterase, by incorporating two bi-functional hemicellulases. Results reveal that non-structured, free hemicellulases out perform those in a cellulosome-like complex in the hydrolysis of wheat arabinoxylan. Moreover, addition of cellulases with xylanosomes slightly improved the release of reducing sugars from destarched corn bran, but not ferulic acid. Overall, the expansion of designer xylanosomes from two to three enzymes was not advantageous in hydrolysis of xylan or gaining access to cellulose. The absence of synergy may be related to restrictive nature of the scaffolding protein, causing structured enzymes to have limited mobility, and therefore, access to their target substrates.

5.2 Introduction

Designer cellulosomes offer significant improvement in cellulose hydrolysis compared to their corresponding free enzyme mixtures [1-3]. However, at the same molar concentration, the rate of cellulose hydrolysis using cellulosomes isolated from *Clostridium cellulolyticum* is considerably higher than that of the most efficient designer cellulosome, resulting almost 9-fold greater release of sugars from Avicel and 2.5-fold from bacterial cellulose [1]. This disparity is most likely due to more enzymes being included in an isolated native cellulosome (minimum of nine in a basic cellulosome of *C. thermocellum* using *CipA*) compared to an *in vitro* cellulosome, which is currently a maximum of four.

Compared to two-unit xylanosomes, the majority of cellulase combinations placed in designer cellulosomes show greater synergy on their respective substrates, reaching synergy ratios of 11 with combinations of GH family 5 and 48 cellulases [1]. Development of tri-functional designer cellulosomes, where 2:1 ratios of endoglucanases to exoglucanases and vice versa showed improvements in cellulose hydrolysis compared to their equivalent bi-functional cellulosomes and free enzyme systems [2]. Using the example of tri-functional cellulosomes, inclusion of an additional hemicellulase within the xylanosome might produce more synergy than the two-unit xylanosome on biomass substrates.

Nature, as well as clever scientists, has provided evidence on the advantages of combining multiple enzyme activities for biomass hydrolysis. Chimeric proteins have been successfully developed that contain three different hemicelluase activities for better arabinoxylan degradation than combinations of the individual enzymes [4]. The presence

of xylanase, arabinofuranosidase, acetylxylan esterase, glucuronidase, ferulic acid esterase, and beta-xylsoidase was observed in the extracellular fraction of *Neosartoya spinosa* NRRL 185 for complete release of ferulic acid from corn bran [5]. However, designer xylanosomes discussed in Chapter 3 only contain two enzymes with a maximum of three different activities in the case of structures containing a bi-functional arabinofuranosidase/xylosidase. As mentioned previously, six different enzyme activities are required to completely hydrolyze xylan [6]. Thus, two-unit xylanosomes are at a disadvantage in completely hydrolyzing the target substrate. Incorporation of more multi-functional enzymes within the xylanosome would increase the overall number of enzymatic activities available within the biocatalyst. Therefore, extended xylanosomes would have the capacity to complete more destruction of the xylan backbone and its side-chain substituents, possibly resulting in greater hydrolysis of xylan.

Chapter 4 describes the characterization of a bi-functional ferulic acid esterase/acetylxylan esterase from *Cellvibrio japonicus*. This enzyme showed synergy with family 10 and family 11 xylanases used in the two-unit xylanosomes in the release of reducing sugars from wheat arabinoxylan, as well as ferulic acid from corn bran. With its multiple activities and observed synergies, Fee1B is a model enzyme to incorporate into xylanosomes, expanding them to contain three components and ultimately making them penti-functional. This chapter discusses the molecular design, construction, and application of two three-unit xylanosomes for the hydrolysis of different biomass substrates.

5.3 Results

5.3.1 Cloning and expression of three-unit xylanosome components

DNA sequences for the *Clostridium cellulolyticum* cohesin and dockerin domains were optimized for *E. coli* expression and artificially synthesized by GenScript, Inc. The truncated version (removal of two carbohydrate binding-domains) of Fee1B from *Cellvibrio japonicus* was fused with the dockerin domain from *CelA* of *C. cellulolyticum* at the C-terminal. The chimeric gene was cloned and expressed into *E. coli* BL21(DE3). The first type I cohesin domain from *CipC* of *C. cellulolyticum* was added to the C-terminal of the SP2 scaffolding protein to generate the three-cohesin SP3 scaffolding protein and successfully expressed using *E. coli* JM109. Each component was purified using nickel-affinity chromatography via histidine tags, desalted, and used for surface plasmon resonance and hydrolysis experiments. Purification results are provided in Table 5.1.

Table 5.1 Purification table for Fee1B-Doc and SP3

Protein	Activity (mU)	Protein (mg)	Specific Activity (mU/mg)	Purification Factor	% Recovery
Fee1B-Doc	4.43	0.77	5.72	2.1	27
SP3	n/a	3.6	n/a	n/a	28

5.3.2 Construction of three-unit xylanosomes

The design of three unit xylanosomes is shown in Figure 5.1. An additional cohesin domain from *C. cellulolyticum CipA* gene was fused to the C-terminal of the existing SP2 protein to generate SP3. The two family 2 and family 35 carbohydrate-binding modules were removed from Fee1B to eliminate possible restriction of xylanosomes due to multiple CBMs, since X10Doc already contains a family 22 CBM that binds to xylan. A dockerin domain from *C. cellulolyticum CelA* was added to the C-terminal the truncated *Fee1B* gene to create F1BDoc.

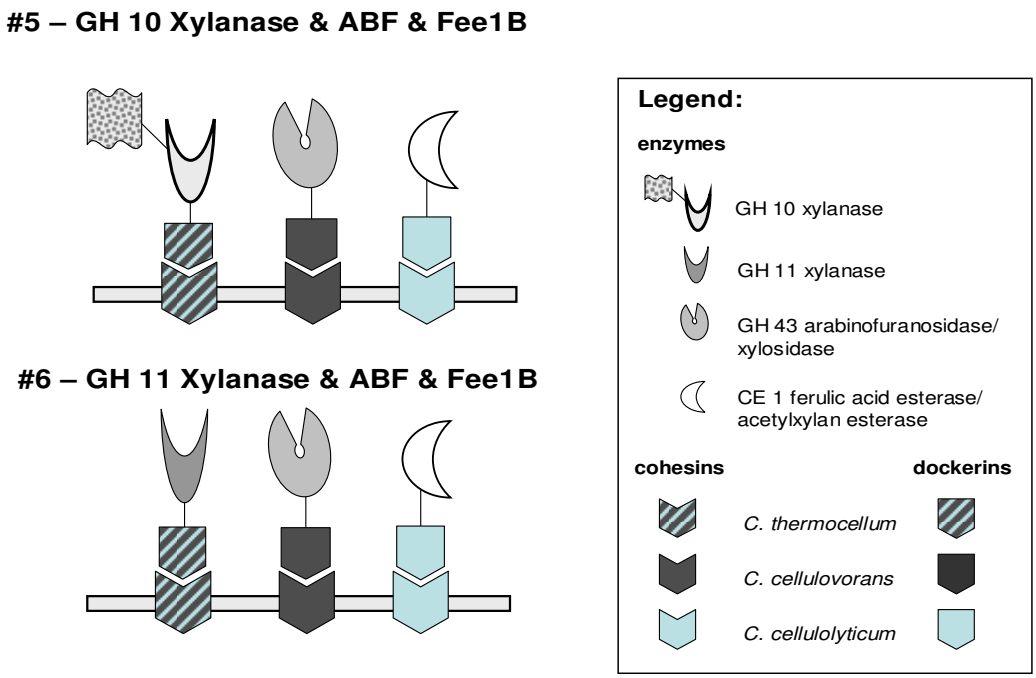


Figure 5.1 Design of three-unit xylanosomes.

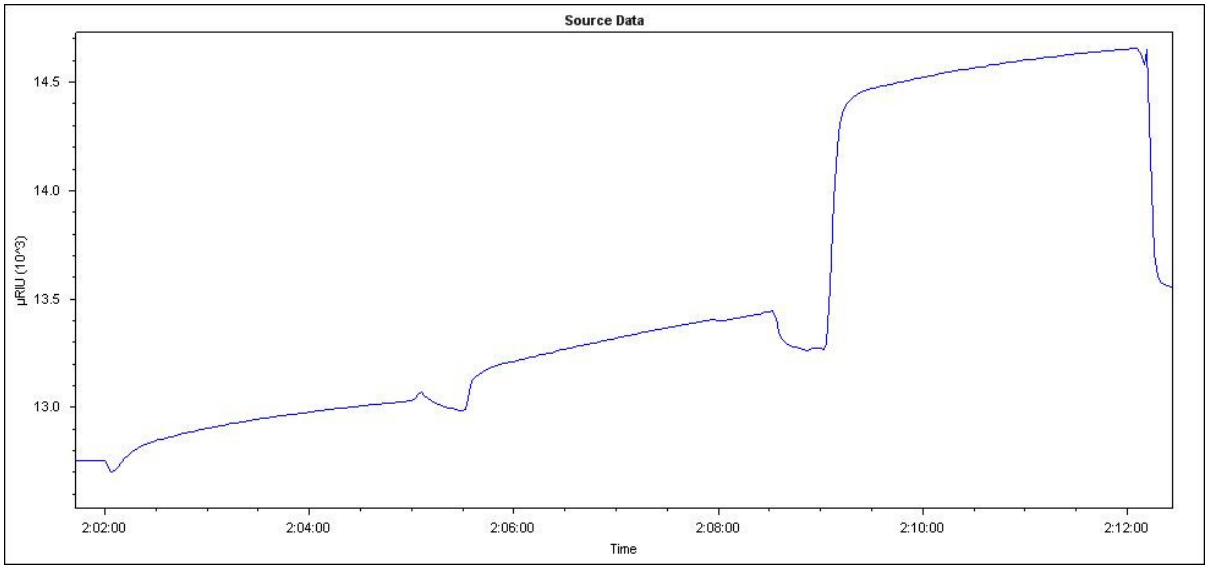
Surface plasmon resonance was used to confirm the proper construction of the three-unit xylanosomes by immobilizing SP3 by its lysine residues. First, the interaction of F1BDoc with SP3 was confirmed. The affinity constant for the *C. cellulolyticum* system was measured and compared well with values reported in literature, as listed in Table 5.2.

Table 5.2 Affinity constant for *Clostridium cellulolyticum* cohesin/dockerin system [7, 8]

	k_{on}	k_{off}	K_A	Literature K_A
C. cellulolyticum	2.2×10^{5} s^{-1}	2.1×10^{-3} (M/s)	1.1×10^{8} (M^{-1})	10^{7} - 10^{10} (M^{-1})

Three-unit xylanosomes were constructed via SPR by sequential addition of dockerin-tagged enzymes. First, F1BDoc was added, followed by ABFDoc, and finally either X10Doc or X11Doc, representing the *C. cellulolyticum*, *C. cellulovorans*, and *C. thermocellum* cohesin/dockerin systems, respectively. Spectra representing these constructions are presented in Figure 5.2. Since xylanosomes were constructed using a sequential addition of F1BDoc → ABFDoc → X10/X11Doc, the same order was used in constructing xylanosomes for biomass hydrolysis reactions.

F1BDoc → ABFDoc → X10Doc



F1BDoc → ABFDoc → X11Doc

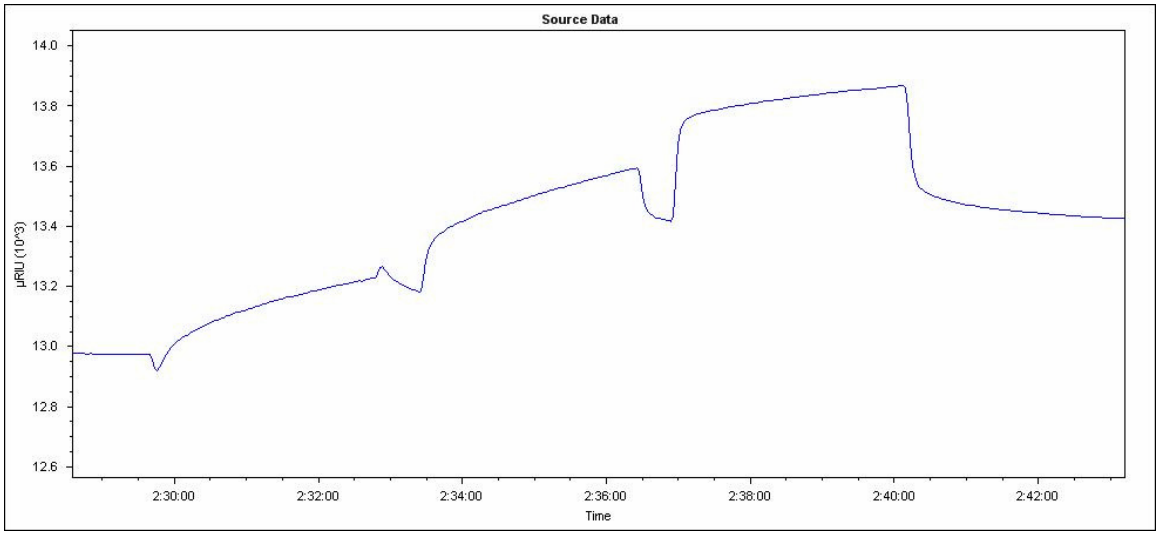


Figure 5.2 SPR spectra of three-unit xylanosome construction.

5.3.3 Biomass hydrolysis using three-unit xylanosomes

5.3.3.1 Release of reducing sugars and ferulic acid from wheat arabinoxylan

Combinations of free hemicellulases and three-unit xylanosomes were used to hydrolyze wheat arabinoxylan (WAX). Unexpectedly, Figure 5.3 reveals that structured hemicellulases perform equal to or lower than the corresponding free enzyme systems in the release of sugars from WAX. While the concentration of released sugars increase, synergy observed on WAX with two-unit xylanosomes discussed in Chapter 3 is not matched nor surpassed with the addition of a third enzyme into the complex.

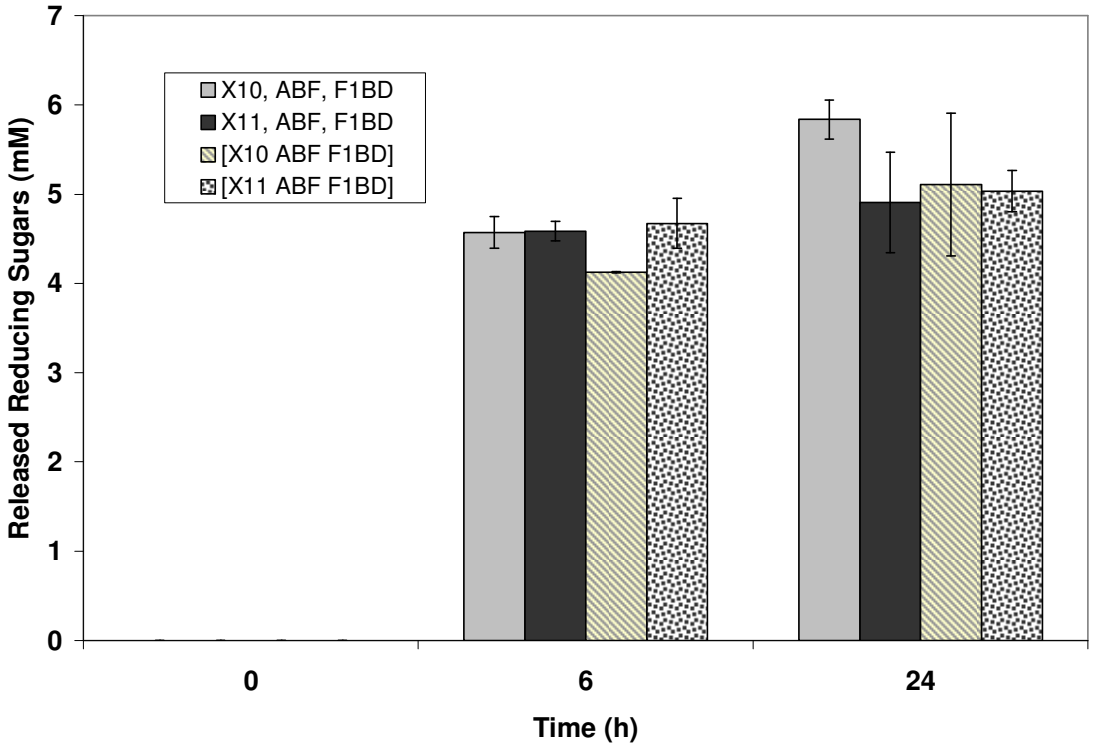


Figure 5.3 Hydrolysis of 1% (w/v) wheat arabinoxylan using free and structured hemicellulases. Results are the average of three independent experiments. Error bars represent standard deviation.

Figure 5.4 compares the release of reducing sugars from WAX using free and structured enzymes in xylanosomes. Out of all combinations, free mixtures of X10Doc, ABFDoc, and F1BDoc outperform both two- and three-unit xylanosomes, suggesting that placing more than three hemicellulases within a scaffolding protein is not beneficial to improving hydrolysis of xylan. In fact, both two- and three-unit xylanosomes behave similarly at by the end of the reaction. However, by adding F1BDoc to the reaction, approximately 20% more reducing sugars was released than with just X10Doc with ABFDoc, giving more evidence on the synergistic effects of adding multiple hemicellulases for xylan hydrolysis.

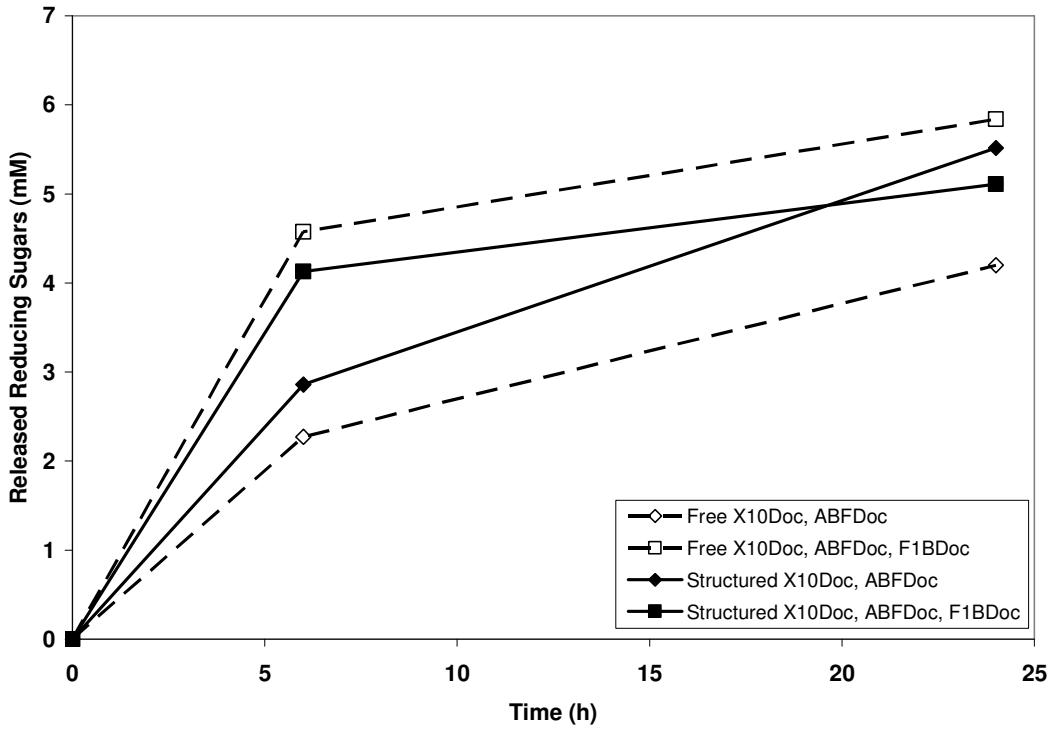


Figure 5.4 Comparison of two- and three-unit xylanosomes in the hydrolysis of 1% (w/v) wheat arabinoxylan. Results are the average of three independent experiments. Error is between 10 – 20% the average value.

Similarly, ferulic acid was not synergistically released from wheat arabinoxylan using three-unit xylanosomes as listed in Table 5.3. This result was not completely unexpected as no synergy was observed when either X10Doc or X11Doc was combined with Fee1B to remove ferulic acid from WAX as discussed in Chapter 4. F1BDoc alone released more ferulic acid, 15 μM, than any of the enzyme combinations, including free systems.

Table 5.3 Release of ferulic acid from wheat arabinoxylan after 24 hours using free and structured hemicellulases[a]

Enzyme Combination	**Ferulic Acid (μM)**	**Synergy**
F1BDoc (free)	15 ± 0.3	-
X10Doc, ABFDoc, F1BDoc (free)	17 ± 0.3	-
X11Doc, ABFDoc, F1BDoc (free)	18 ± 0.4	-
[X10Doc + ABFDoc + F1BDoc]	18 ± 0.2	1.05
[X11Doc + ABFDoc + F1BDoc]	19 ± 0.5	1.07

[a] Synergy is calculated dividing the product released of enzyme mixtures by the sum of product released by each individual enzyme.

The application of three-unit xylanosomes containing three hemicellulases, with two of them possessing multiple activities, does not improve release of sugars or ferulic acid from WAX. However, based on previous results with xylanosomes on destarched corn bran, the three-unit structures were evaluated for their ability to expose cellulose fibrils for increased cellulase activity.

5.3.3.2 Release of reducing sugars and ferulic acid from destarched corn bran

Three-unit xylanosomes were applied to destarched corn bran (DCB) in the presence an endoglucanase and exoglucanase. Undetectable levels of reducing sugars were released from DCB by free or structured hemicellulases alone; cellulases were required in order to perform the reducing sugar assay. Figure 5.5 shows how xylanosomes performed with cellulases in the release of reducing sugars from DCB.

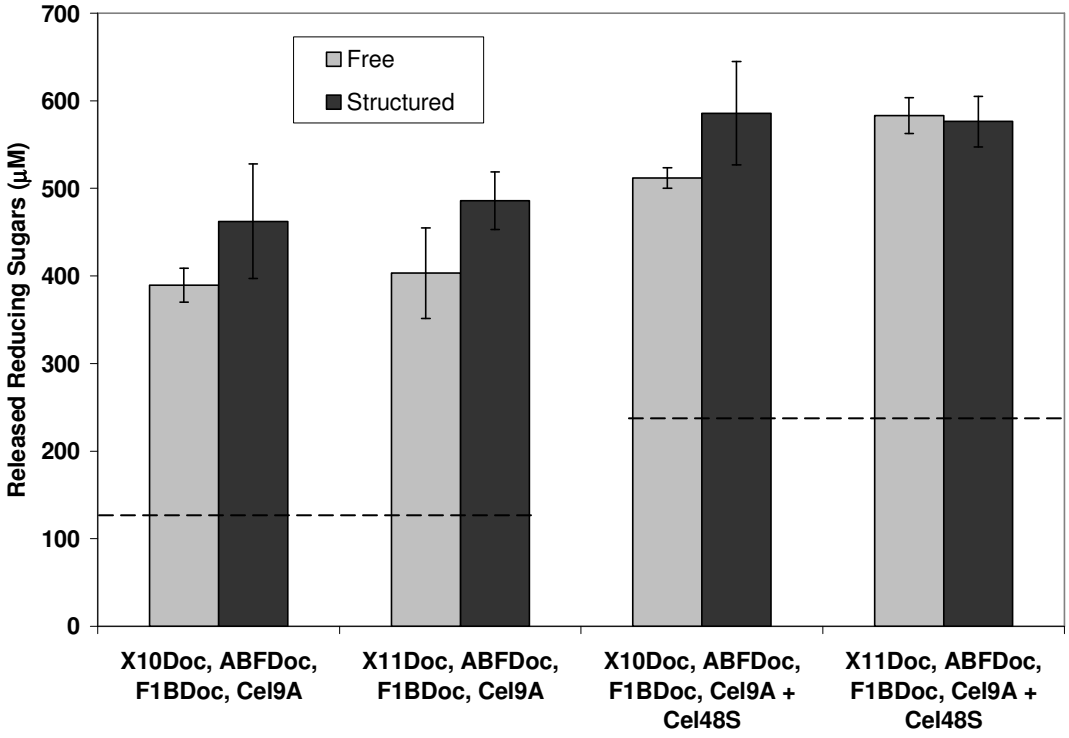


Figure 5.5 Hydrolysis of 1% (w/v) destarched corn bran using free and structured hemicellulases with free cellulases. Results are the average of three independent experiments. Error bars represent standard deviation. Cel9A and combinations of Cel9A + Cel48S release 122 ± 35 μM and 235 ± 12 μM reducing sugars from DCB, respectively (dotted lines). Cel48S alone released no detectible sugars.

With addition of the endoglucanase Cel9A, three-unit xylanosomes allow for greater degradation of DCB, releasing 20% more sugars compared to free enzyme systems. This result most likely due to increased exposure of the cellulose for Cel9A by better removal of hemicellulose (xylan) with structures versus free systems. Further inclusion of Cel48S in the hydrolysis reaction gave an even higher release of total sugars than all reactions containing Cel9A. Three-unit structures with X10Doc gave 74 mM more sugars that free systems, whereas those structures with X11Doc resulted in nearly the same concentrations, which may be due to the different modes of action of the two xylanases. Synergies of structures with Cel48S plus Cel9A were lower than those with Cel9A alone. Unlike the two-unit xylanosomes on DCB, three-unit structures do not provide a significant advantage over the free enzyme systems in sugar release, signifying the addition of another enzyme to the xylanosome is ineffective.

Figure 5.6 compares two- and three-unit xylanosomes in the release of reducing sugars from DCB. Data reveal all three-unit xylanosomes provide better hydrolysis of DCB compared to two-unit structures, reiterating that addition of F1BDoc was significant in improving hydrolysis of the complex substrate. Interestingly, combining two-unit xylanosomes (X11Doc, ABFDoc) with Cel9A accomplish better hydrolysis than free X11Doc and ABFDoc with both Cel9A and Cel48S. But overall, two-unit structures provided more synergy than their three-unit counterparts.

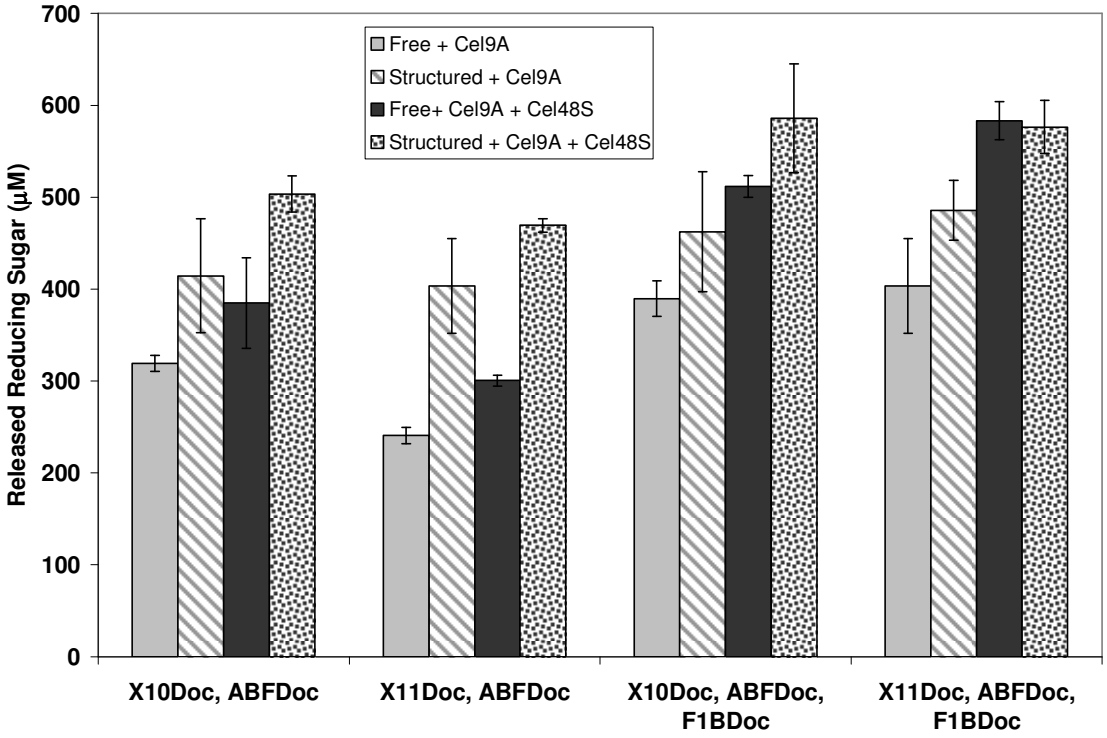


Figure 5.6 Comparison of two- and three-unit xylanosomes on destarched corn bran. Results are the average of three independent experiments. Error bars represent standard deviation. Cel9A and combinations of Cel9A + Cel48S release 122 ± 35 μM and 235 ± 12 μM reducing sugars from DCB, respectively. Cel48S released no detectible sugars.

On the other hand, no synergy was observed in the removal of ferulic acid with destarched corn bran as listed in Table 5.4. While addition of F1BDoc was critical in improving release of reducing sugars from DCB, it appears that its action on ester-linked ferulic acid is hindered. Therefore, the more important enzymatic activity in hydrolysis is perhaps the acetylxylan esterase; removal of acetyl groups may be critical in eliminating hydrogen bonding between cellulose and xylan polymers, opening access for cellulases and hemicellulases to act.

Table 5.4 Release of ferulic acid from destarched corn bran after 24 hours using cellulases and hemicellulases[c,d]

Enzyme Combination	Ferulic Acid (μM)			Synergy	
	+ *Cel9A*	+*Cel48S*	+ *Cel9A, Cel48S*	+ *Cel9A*	+ *Cel9A, Cel48S*
X10Doc, ABFDoc, F1BDoc (free)	26 ± 0.9	23 ± 1.4	30 ± 1.0	-	-
X11Doc, ABFDoc, F1BDoc (free)	28 ± 1.1	23 ± 0.5	39 ± 1.3	-	-
[X10Doc + ABFDoc + F1BDoc]	22 ± 0.5	20 ± 0.4	24 ± 0.2	0.86	0.82
[X11Doc + ABFDoc + F1BDoc]	20 ± 0.2	19 ± 0.4	23 ± 0.1	0.73	0.59

[c] F1BDoc released 19 ± 0.7 μM ferulic acid from DCB.
[d] Synergy is calculated dividing the product released of enzyme mixtures by the sum of product released by each individual enzyme.

5.4 Discussion

Designer xylanosomes containing three different hemicellulases were constructed and used to release reducing sugars and ferulic acid from two biomass substrates containing xylan: wheat arabinoxylan and destarched corn bran. Cohesin/dockerin domain systems from *C. thermocellum*, *C. cellulovorans*, and *C. cellulolyticum* allowed for formation of a biocatalyst with known composition of either GH family 10 or family 11 xylanase, bi-functional arabinofuranosidase/xylosidase, and bi-functional ferulic acid esterase/acetylxylan esterase, respectively. A functional *C. cellulolyticum* dockerin domain was successfully appended to the C-terminal of a truncated version of *Fee1B* from *C. japonicus*, further demonstrating the feasibility of incorporating non-cellulosomal enzymes into cellulosomal-like structures.

Compared to two-unit xylanosomes, these structures have less activity on wheat arabinoxylan (WAX), generating synergy ratios of less than 1 in both the release of reducing sugars and ferulic acid. Thus, it appears that for a hemicellulose-rich substrate, such as WAX, placing more than two enzymes within a xylanosome results in hindering the accessibility of the bound enzyme to its target sites within the biomass. Another possible explanation for the loss of activity is the overall structure of WAX; the frequency of side-chain substitutions (i.e., ferulic acid, arabinose) may not match the spacing of the enzymes within the three-unit xylanosome, causing it to be less effective. For example, native cellulosomes combine nine different enzymes into one large MDa complex and the same activity, such as endo- and exo-glucanases, may be present multiple times to cover sizable areas of the cellulose substrate, allowing the overall structures to be more effective [9]. In the instance of two-unit xylanosomes, hydrolysis

could be easier for this reasoning; those containing a xylanase with ferulic acid esterase may preferentially act on areas of the backbone with ferulic acid side chains. Similarly, structures containing xylanase and an arabinofuranosidase are active on more areas where arabinose is present. For three-unit xylanosomes, it would be expected that since ferulic acid is found ester-linked to arabinose, that these structures would show greater synergy on WAX. However, mobility of enzymes to properly interact with their target substrate is more sterically hindered with the addition of a third enzyme within the scaffolding protein.

Other biomass substrates which contain cellulose as well as hemicellulose allow xylanosomes to have an additional role in hydrolysis. During the hydrolysis of xylan on corn bran, it is assumed that cellulose fibrils embedded within a hemicellulose network become more accessible, increasing the number of sites on which cellulases can be active. This theory is supported by the synergy observed when three-unit xylanosomes are added with a GH family 9 endoglucanase or a combination of endoglucanase with a GH family 48 cellobiohydrolase for hydrolysis on destarched corn bran (DCB). By using a xylanosome, it is likely larger areas of hemicellulose within DCB are removed, instead of the possible random cleavages given by free enzymes. Thus, cellulases can gain greater access to cellulose and in return, release more reducing sugars from the biomass. Conversely, three-unit xylanosomes did not act synergistically release ferulic acid from DCB, continuing the trend seen with two-unit xylanosomes containing FAEDoc. Thus, neither DCB nor WAX are substrates on which three-unit xylanosomes can be effectively applied for ferulic acid recovery, despite that corn residues contain a high concentration of ferulic acid of around 3% (w/w).

In conclusion, three-unit xylanosomes are effective at opening up the biomass structure by removing hemicellulose to expose cellulose for further degradation. While they do not synergistically hydrolyze xylan substrates, they can still play a valuable role in improving enzymatic hydrolysis for biomass utilization.

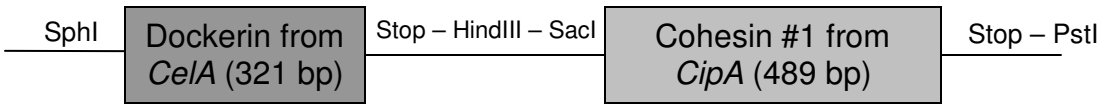
5.5 Materials and Methods

5.5.1 Strains and materials

Strains for X10Doc, X11Doc and ABFDoc are described in Chapter 2. DNA sequences encoding cohesin and dockerin domains from *Clostridium cellulolyticum* were prepared by GenScript USA Inc. and cloned into the general cloning vector, pUC57. Genomic DNA of *Cellvibrio japonicus* was prepared as described in Chapter 4. Wheat arabinoxylan (Megazyme, Ireland) and destarched corn bran (Sun Opta, Bedford, Massachusetts, U.S.A) were used as biomass substrates. Destarched corn bran was prepared as described in Chapter 3.

5.5.2 Cloning and expression of F1BDoc and SP3

The cloning vector containing the sequences for both *C. cellulolyticum* cohesin and dockerin domains was named pCcelCD and transformed into *E. coli* JM109 for vector propagation. Linker sequences before (TVLPKDIPGDS) and after (TIDPGTQPTKE) the cohesin domain, as well as prior to the dockerin domain (GELFGFFRRS), were included in the synthesized gene. Sequence orientation is shown in Figure 5.7.



Total Gene Size – 810 bp

Figure 5.7 Schematic of synthesized *C. cellulolyticum* cohesin/dockerin system with denoted restriction sites.

Primers designed for cloning of each domain are listed in Table 5.5. The truncated version of Fee1B was amplified from the genomic DNA of *C. japonicus*, fused to the dockerin of *C. cellulolyticum* using *Sal*I restriction sites, and inserted into *Bam*HI-*Hind*III linearized pQE80L to produce pQF1BDoc. SP2 was amplified from pQSP2 and digested with *Bgl*II and *Sac*I. The cohesin domain of *C. cellulolyticum* was added to the C-terminal of SP2 using *Sac*I and *Hind*III to give pQSP3. *E. coli* JM109 was transformed with either recombinant plasmid pQF1BDoc or pQSP3 and used to express F1BDoc and SP3, respectively.

Table 5.5 Primer sets used in the construction of F1BDoc and SP3

Target Gene	**Primer Set (5'→3')**
C. cellulovorans Dockerin (pCcelCD)	CcelD – F: ATCCGATGACCCAATGCA CcelD – R: CAGGACGGTCAGTTCAAGCT
Truncated Fee1B (*C. japonicus*)	XFee1B – F: AAAGAAGGATCCGGTAAACCGCGCACAGTG XFee1B – R: AGTAGTGCATGCGAACTGTGTGAAGAATTGCCAGG
C. cellulovorans Cohesin (pCcelCD)	CcelC – F: TAAAAAGTGGAGCTCACCGTC CcelC – R: TAGATTTGGGTCCTGCAGTCAT
SP2 (removal of stop codon, pQSP2)	SP2c – F: CTACTAAGATCTGGTGGCGCAGCTATGATACC SP2c – R: CCTTTGAGCTCGCTAACTTTCACACTTCCGTTAACTGT

F1BDoc was expressed using 0.5 mM IPTG at 25^{o}C for 20 hours in LB media containing 100 μg/mL ampicillin. Similarly, SP3 was expressed using 0.5 mM IPTG at 30^{o}C for 5 hours. Crude lysates were prepared as described in Chapter 2 and each protein purified using nickel-affinity chromatography via N-terminal histidine tags present on each protein.

5.5.3 Surface plasmon resonance

SP3 was immobilized onto a SR700 gold sensor slide with a mixed self-assembled monolayer of 90% OH-$(PEG)_6$-C_{11}-SH/10% COOH-$(PEG)_6$-C_{11}-SH. The slide was functionalized using 0.2M N-(3-Dimethylaminopropyl)-N'-ethylcarbodiimide (EDC, Sigma-Aldrich, U.S.A.) and N-hydroxysuccinimide (NHS, Sigma-Aldrich,

U.S.A.), then exposed to SP2 (20 μg/mL) in 10 mM sodium acetate, pH 4.4, followed by 1M ethanolamine to cap any unreacted sites. All experiments were performed in 20 mM Tris-maleate, 1 mM CaCl2, 0.05% (v/v) Tween 20, pH 6.0, except SP3 immobilization where the running buffer was Phosphate-buffered Saline (Bio-Rad) with 0.05% (v/v) Tween 20 (Sigma-Aldrich). F1BDoc, ABFDoc, X10Doc, and X11Doc were all dialyzed into the same buffer. The affinity constant for F1BDoc was determined using protein concentrations of 10 nM, 25 nM, and 50 nM using the method described in Chapter 2, section 2.3. Xylanosomes were constructed using 50 nM concentrations of the dockerin-tagged enzymes.

5.5.4 Xylanosome construction and biomass hydrolysis reactions

Xylanosome construction was performed by adding equimolar quantities of appropriate dockerin-tagged enzymes in 20 mM Tris-maleate, 1 mM $CaCl_2$, pH 6.0 at room temperature in 1.5-mL microcentrifuge tubes. Equimolar quantities of the scaffolding protein, SP3, were added to the solutions of F1BDoc, followed by ABFDoc, and lastly, either X10Doc or X11Doc for structure formation. Gentle vortexing and brief centrifugation (less than 3 seconds) was performed after adding each dockerin-tagged enzyme. Proteins were added to obtain a final concentration of 0.1 μM for a 1 mL reaction volume.

E. coli E609Y/pQpelCelA and *E. coli* E609Y/pQTHCelS were cultured in LB with ampicillin (100 μg/mL) and induced with 0.5 mM IPTG for 20 hours at 22^{o}C for extracellular expression of Cel9A and Cel48S. Cells were removed via centrifugation (4000 x g, 25 minutes, 4^{o}C). Resulting supernatants were concentrated three times using dialysis tubing (MW 10,000 Da cutoff) and polyethylene glycol. Cellulase activity was

measured using 1% (w/v) carboxymethyl cellulose (CMC, 30 minutes, 40^{o}C) and reducing sugars determined as described by Miller (1959). Cel5A activity was 1024 U/mL concentrated supernatant. Cel48S had no detectible activity on CMC.

For hydrolysis of biomass substrates, 5% (w/v) wheat arabinoxylan or 6% (w/v) destarched corn bran in 20 mM Tris-maleate, 1 mM $CaCl_2$, pH 6.0, were added to xylanosomes and incubated at 50^{o}C for 24 hours at 1000 rpm in a thermoixer. In reactions with cellulases, 150 U (or 150 µL) of Cel9A and 150 µL of Cel48S were added to the reaction. To stop the hydrolysis reaction, mixtures were exposed to boiling water for 5 minutes. After cooling, reaction mixtures were centrifuged for 10 minutes at 16,000 x *g* to pellet insoluble biomass and the supernatant removed for subsequent analysis.

5.5.5 Analytical methods

Protein concentration of xylanosome components was measured by the method described by Bradford using a protein reagent dye (Bio-Rad), with bovine serine albumin as the standard. To measure released reducing sugars, appropriately diluted supernatants were combined with dinitrosalicyclic (DNS) acid solution as outlined in Miller [10] and incubated in boiling water for 5 minutes for the colorimetric reaction to occur. After cooling, the absorbance at 550 nm was measured and compared to a calibration curve generated with xylose as the standard. Quantification of released ferulic acid was performed on hydrolysis supernatants by high performance liquid chromatography as described by Shin and Chen [11].

5.6 References

1. Fierobe, H.P., et al., *Degradation of cellulose substrates by cellulosome chimeras. Substrate targeting versus proximity of enzyme components.* J Biol Chem, 2002. **277**(51): p. 49621-30.

2. Fierobe, H.P., et al., *Action of designer cellulosomes on homogeneous versus complex substrates: controlled incorporation of three distinct enzymes into a defined trifunctional scaffoldin.* J Biol Chem, 2005. **280**(16): p. 16325-34.

3. Mingardon, F., et al., *Incorporation of fungal cellulases in bacterial minicellulosomes yields viable, synergistically acting cellulolytic complexes.* Appl Environ Microbiol, 2007. **73**(12): p. 3822-32.

4. Fan, Z., et al., *Multimeric hemicellulases facilitate biomass conversion.* Appl Environ Microbiol, 2009. **75**(6): p. 1754-7.

5. Shin, H.D., et al., *A complete enzymatic recovery of ferulic acid from corn residues with extracellular enzymes from Neosartorya spinosa NRRL185.* Biotechnol Bioeng, 2006. **95**(6): p. 1108-15.

6. Saha, B.C., et al., *Lignocellulose biodegradation*. ACS symposium series 889. 2004, Washington, DC: American Chemical Society : Distributed by Oxford University Press. xii, 400 p.

7. Fierobe, H.P., et al., *Cellulosome from Clostridium cellulolyticum: molecular study of the Dockerin/Cohesin interaction.* Biochemistry, 1999. **38**(39): p. 12822-32.

8. Pinheiro, B.A., et al., *The Clostridium cellulolyticum dockerin displays a dual binding mode for its cohesin partner.* J Biol Chem, 2008. **283**(26): p. 18422-30.

9. Raman, B., et al., *Impact of pretreated Switchgrass and biomass carbohydrates on Clostridium thermocellum ATCC 27405 cellulosome composition: a quantitative proteomic analysis.* PLoS One, 2009. **4**(4): p. e5271.

10. Miller, G.L., *Use of dinitrosaliclic acid reagen for determination of reducing sugars.* Analytical Chemistry, 1959(31): p. 127-132.

11. Shin, H.D. and R.R.Z. Chen, *Production and characterization of a type B feruloyl esterase from Fusarium proliferatum NRRL 26517.* Enzyme and Microbial Technology, 2006. **38**(3-4): p. 478-485.

CHAPTER 6

CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE DIRECTIONS

6.1 Conclusions

The work presented in this dissertation accomplished the three objectives stated in the introduction: 1) six different xylanosomes were designed and constructed, 2) the protein-protein interaction between cohesins and dockerins was analyzed at various conditions, and 3) xylanosomes were used to improve hydrolysis of wheat arabinoxylan and destarched corn bran over the corresponding free enzyme systems.

6.1.1 Molecular design and construction of xylanosomes (Chapter 2 and Chapter 5)

The design and construction of the first cellulosome-like structures specifically designed for xylan (hemicellulose) degradation was achieved. Five hemicellulases and two chimeric scaffolding proteins were cloned and expressed into *Escherichia coli* for construction of four 2-unit xylanosomes and two 3-unit xylanosomes. Induction conditions using IPTG were optimized and xylanosome components were produced in quantities ranging from 0.3 – 22 mg/L culture. A total of seven proteins were successfully purified using immobilized metal (nickel) affinity chromatography using either N-terminal (X10Doc, X11Doc, F1BDoc, SP2, SP3) or C-terminal (FAEDoc, ABFDoc) histidine tags. Construction of all six xylanosomes was performed using surface plasmon resonance. Both scaffolding proteins, SP2 and SP3, were individually immobilized on a functionalized gold-plated glass slide via lysine residues. Spectra

revealed that some two-unit xylanosomes required sequential addition of enzymes to the scaffolding protein, SP2, revealing that steric hindrances can inhibit proper formation. Both three-unit xylanosomes were successfully constructed using a sequential addition of enzymes. The order of enzyme addition for structure formation from SPR analysis was used to construct xylanosomes for hydrolysis reactions. While designer cellulosomes do not require such conditions for construction with cellulosomal enzymes [1-6], non-cellulosomal hemicellulases seem to behave differently and need such procedures for incorporation into these types of structures.

6.1.2 Characterization of self-assembly via cohesins and dockerins (Chapter 2, Chapter 3)

The cohesins and dockerins of cellulosome-producing microbes possess a strong, calcium-dependent, high-affinity interaction, which ultimately allows cellulosomes to form extracellularly. Surface plasmon resonance was used to measure the affinity constants for the cohesin/dockerin systems used from *Clostridium thermocellum* and *Clostridium cellulovorans*. This interaction was characterized at various temperatures, as well as in the presence of ethanol and non-specific proteins, to mimic possible consolidated bioprocessing conditions. Dockerins within chimeric enzymes used for xylanosome construction gave affinity constants that compare well to those in literature, proving that non-cellulosomal hemicellulases can be incorporated into cellulosomal structures. Analysis of the *C. thermocellum* cohesin and dockerin system revealed that the domains maintain a high affinity interaction in the presence of high concentrations of non-dockerin or non-cohesin bearing proteins. Thus, xylanosomes, as well as other cellulosome-like structures, can be formed without initial purification of the components.

This attribute has recently been used in formation of designer cellulosomes on the surface of *Saccharomyces cerevisiae* [7, 8].

Xylan hydrolysis reactions were performed at elevated temperatures (50^{o}C) compared to those used for cellulose hydrolysis. Therefore, affinity constants of the thermophillic *C. thermocellum* and mesophilic *C. cellulovorans* cohesin/dockerin systems were measured at various temperatures. Results confirm the slightly exothermic nature of dockerins binding to cohesins [9], with affinity constants increasing in magnitude with increasing temperature. While consolidated bioprocessing conditions may or may not exceed temperatures of 60^{o}C in the future, it is evident that xylanosomes will be able to form at currently used enzymatic saccharification and microbial fermentation temperatures.

Lastly, binding of *C. thermocellum* and *C. cellulovorans* dockerins was investigated in the presence of ethanol to learn how xylanosomes can form during fermentation of sugars. Affinity constants were found to decrease two orders of magnitude in the presence of 10% (v/v) ethanol for both systems. The cohesin-dockerin interaction was reported as being hydrophobic [9] and the addition of ethanol to water changes the hydrophobicity of the buffer solution. This phenomenon could cause the necessary serine and threonine amino acids of dockerins [10] to interact more frequently with the surrounding solution than with the cohesin domains or even slightly alter the overall structure of the cohesin and dockerin domains, making binding less optimal. However, ethanol concentrations of 10% (v/v) are on the higher end those seen from microbial fermentations, especially from five-carbon sugars, such as xylose. Ideally, the enzymatic hydrolysis of the biomass substrate has come to completion, eliminating the

need for xylanosomes to maintain their structure, or the catalytic activity of the enzymes has been lost.

Overall, with the studies performed, cohesins and dockerins maintain their high-affinity interaction with high temperature, varied concentrations of non-specific proteins, and in the presence of ethanol. These attributes make xylanosomes very attractive for use in consolidated bioprocessing applications for converting biomass to biofuels and other commodity or value-added chemicals.

6.1.3 Characterization of the Fee1B from *Cellvibrio japonicus* and its role in xylan hydrolysis (Chapter 4)

The recent genome sequencing of *Cellvibrio japonicas* [11] revealed the presence of two highly homologous ferulic acid esterases (FAEs), encoded by fee1A and fee1B. The previously unknown FAE, Fee1B, was successfully cloned and expressed in an *E. coli* system and characterized as a type-D FAE based on substrate specificity on methyl esters of hydroxycinnamic acids. Furthermore, it showed activity on *p*-nitrophenyl acetate, making it also an acetyl xylan esterase. Fee1B gave a pH and temperature optima of 6.5 and between 35-40^{o}C, respectively. The pH stability of the enzyme was maintained over a wide range of pH. Additionally, the multi-domain enzyme is highly unusual in its presence of two tandem N-terminal carbohydrate binding modules (CBMs) belonging to family 2 and 35. These CBMs were shown to be crucial for optimum enzyme activity on synthetic substrates. The potential of the enzyme in biomass processing was demonstrated with its high synergy with either a GH family 10 or 11 xylanase in the release of reducing sugar from arabinoxylan and its ability to liberate ferulic acid from wheat arabinoxylan and synergistically from corn bran. The bi-

functional activity of Fee1B, as well as its propensity to enhance xylan hydrolysis with xylanases already used in two-unit xylanosomes, made this enzyme an ideal addition in the design and construction of three-unit xylanosomes.

6.1.4 Application of xylanosomes for biomass utilization (Chapter 3, Chapter 5)

Successful hydrolysis of wheat arabinoxylan and corn bran was observed with xylanosomes. Sugars were released between 15 to 30% more from wheat arabinoxylan using two-unit xylanosomes. Furthermore, two-unit structures also synergistically released ferulic acid from wheat arabinoxylan. Conversely, three-unit xylanosomes were not effective on WAX in either the release of sugars or ferulic acid, signifying that expanding structures to include more enzymatic activities is not always beneficial for biomass hydrolysis. As a result of hydrolyzing xylan more effectively than free enzymes, two-unit xylanosomes were shown to improve activity of an endoglucanase and cellobiohydrolyase on destarched corn bran by increasing accessibility of cellulose within the substrate. While three-unit xylanosome showed some synergy with cellulases on destarched corn bran, it was lower than that observed with the two-unit structures, further signifying that placing more hemicellulases within xylanosomes does not improve hydrolysis. Unfortunately, ferulic acid release was not improved from destarched corn bran with any of the xylanosomes, indicating that substrate selection can also affect the effectiveness of the multi-functional biocatalysts.

The observed synergies using xylanosomes are not as high as those seen with designer cellulosomes. This gap may be related to the mode of enzymatic action on cellulose versus hemicellulose (i.e., method of hydrolysis). Cellulose requires only three enzymes to completely hydrolyze the polymer into glucose, with the product of

endoglucanases being a substrate for cellobiohydrolases, and the product of cellobiohydrolases being the substrate for beta-glucosidases. Therefore, placing these cellulases (endoglucanases, cellobiohydrolases) in a scaffolding-like protein allows for close proximity to not only the cellulose substrate, but each other for a type of "relay" hydrolysis mechanism, where each enzyme is feeding another. Hemicellulose has a higher level of complexity than cellulose. The xylan backbone itself only requires xylanase and beta-xylosidase activities to break it down. However, the backbone is frequently substituted with acetyl groups, sugars, and sugar acids that can hinder xylanase or beta-xylosidase action. Therefore, continued design and composition of xylanosomes is needed to alter the structures to be more effective in xylan hydrolysis.

6.2 Significant Contributions

This dissertation provides significant contributions to the field of enzymatic hydrolysis for improving overall biomass utilization. First, major contributions were made in the development of self-assembling, multi-functional biocatalysts. Work presented describes the first instance of cellulosome-like protein nanostructures specifically targeted for xylan hydrolysis, termed xylanosomes. These xylanosomes contained up to three different dockerin-tagged hemicellulases in scaffolding proteins possessing two or three cohesin domains for formation of bi- or tri-functional protein structures. Three non-cellulosomal enzymes were successfully inserted into xylanosomes: a xylanase from *Bacillus halodurans* [12], a bi-functional arabinofuranosidase/xylosidase from the metagenome of a compost starter mixture [13], and a previously uncharacterized bi-functional acetylxylan esterase/ferulic acid esterase from *Cellvibrio japonicus*. Thus, in addition to cellulases, hemicellulases not found in

cellulosomes can also be tagged with dockerin domains and used in designer xylanosome or cellulosomes. Furthermore, xylanosomes developed in this dissertation provide evidence for the first cellulosome-like designer structures to utilize the *Clostridium cellulolyticum* cohesin and dockerin domain system. Research presented also gave more examples of dockerin domains reducing catalytic activity of enzymes, where both cellulosomal and non-cellulosomal hemicellulases showed higher V_{max} values without the dockerin domain. Michaelis-Menten (K_M) concentrations remained similar regardless of the presence of a dockerin domain.

This dissertation also increased knowledge of how cohesins and dockerins behave at typical consolidated bioprocessing conditions, such as higher temperatures and in the presence of high concentrations of ethanol. The measured affinity constants show retention of strong protein-protein interaction between type I dockerins and cohesins up to 60^{o}C, as well as in 10% (v/v) ethanol. Also, evidence was provided showing high-affinity interaction for cohesin and dockerin domains in the presence of non-specific proteins, eliminating the need for protein purification for xylanosome construction. This information is important in the development of stable and functional cellulosome-like biocatalysts for improving enzymatic hydrolysis of biomass.

Lastly, xylanosomes generated in this dissertation were shown to successfully improve hydrolysis of two hemicellulose-containing substrates, wheat arabinoxylan and destarched corn bran. By improving hydrolysis of hemicellulose, more sugars and other components within biomass can be utilized for producing biofuels and other value-added chemicals. Research performed in this dissertation will most likely be used to further

advance enzymatic hydrolysis of biomass, yielding a more economical option in biomass utilization.

6.3 Recommendations for Future Directions

Three objectives were accomplished in this dissertation: 1) two- and three-unit xylanosomes were designed and constructed, 2) the cohesin-dockerin interaction was characterized at conditions for possible consolidated bioprocessing applications, and 3) xylanosomes were successfully applied for the hydrolysis of biomass substrates. However, modest increases in synergy were observed by placing hemicellulases in structured form versus free systems was most likely due to the heterogeneity of xylan, where there is not a "relay" mechanism for hemicellulases as suggested with cellulases. Future directions for improving xylanosome performance include further development of xylanosomes as individual structures and for whole-cell bio-catalysis.

6.3.1 Revisiting molecular design of xylanosomes

Attempts to improve synergy by expanding xylanosomes from 2-unit to 3-unit, as well as increasing the overall number enzymatic activities include five out of six required to degrade xylan, were only successful on select substrates and with the addition of auxiliary cellulases. As mentioned in the Section 6.1.4, the mechanism of hemicellulose hydrolysis does not considerably lend itself to degradation using xylanosomes as cellulose does with cellulosomes. The modest levels of synergy observed with xylanosomes is likely due to significant differences in how hemicellulose is hydrolyzed compared to cellulose. The homogeneity of cellulose allows for efficient hydrolysis using cellulosomes; the varied structure and substitutions of hemicellulose does not lend itself to the same mechanism. Thus, the mode of synergy between cellulases ("relay")

differs from the same between hemicellulases. Therefore, other approaches in the design and composition of xylanosomes may provide better results compared to those discussed in this dissertation.

6.3.1.1 Use of cellulosomal bi-functional enzymes

In their natural environment, cellulosomes come into contact with complex biomass substrates, unlike the isolated cellulose substrates frequently used in designer cellulosome studies. To handle hemicellulose, cellulosomes contain xylanases as well as other hemicellulases to expose embedded cellulose fibrils. For example, multiple bi-functional hemicellulases are found within the native cellulosomes of *C. thermocellum* and *C. cellulovorans*. The modular *XynY* and *XynZ* genes possess both xylanase and ferulic acid esterase domains [14]. Moreover, *XynA* from *C. cellulovorans* contains a GH family 11 xylanase along with an acetylxylan esterase domain [15]. Selection of cellulosomal bi-functional hemicellulases for use in two- or three-unit scaffoldings could be more advantageous than converting non-cellulosomal enzymes. Because these enzymes are originally found in structured form, the spacing and orientation of the catalytic domains have been optimized for xylan hydrolysis within those structures. However, the length of linker sequences between dockerin domains and non-cellulosomal cellulases was found to have minimal affect on enzyme activity [16]. This may or may not be the case with bi-functional enzymes, especially hemicellulases. Furthermore, the multi-functionality of bi-functional cellulosomal hemicellulases will allow inclusion of more activities for complete hemicellulose degradation without additional cloning or enzyme screening. Xylanosomes with cellulosomal bi-functional hemicellulases might be most beneficial for use with complex biomass substrates for increasing exposure of

cellulose. Additionally, performing reactions with such xylanosomes in conjunction with other designer cellulosomes or non-structured cellulases could significantly improve overall enzymatic hydrolysis of biomass and its utilization.

6.3.1.2 Increased use of backbone-acting enzymes

Another design approach for improving xylanosome activity would be to include multiple xylanases within current or extended scaffolding proteins. High molecular weight protein complexes containing various xylanase activities were identified in nature *Streptomyces olivaceoviridis* [17-19]. Similarly, native cellulosomes are found to possess multiple cellulases with the same or similar activities to efficiently hydrolyze the recalcitrant cellulose substrate [20]. Thus, various combinations of xylanases, both family 10 and family 11, which act on different parts of the xylan backbone, can also be utilized within one xylanosome, mimicking the protein nanostructures found in nature. Furthermore, cellulosomes are significantly larger (1 – 2 MDa) compared to the 200 – 240 kDa of three-unit xylanosomes. Increasing the number of cohesins within the scaffolding protein would provide a biocatalyst that could span larger areas of the xylan backbone, allowing each type of xylanase easier access to their respective target sites. In addition, use of three or more cohesins from different species would allow incorporation of both families of xylanase, as well as a beta-xylosidase. While a monomeric xylosidase was used in this dissertation, the majority of beta-xylosidases are found as multimers [21], so discovery of more suitable proteins would be ideal to further expand the capabilities of such xylanosomes. These backbone-acting xylanosomes could be engineered for generation of smaller xylo-oligosaccharides and monosaccharides for direct fermentation into biofuels or bio-based chemicals.

6.3.1.3 Incorporation of carbohydrate-binding modules

The scaffolding proteins of *C. thermocellum* and *C. cellulovorans* both contain a family 3a carbohydrate-binding module (CBM) for specific binding to crystalline cellulose. CBMs place cellulosomes and enzymes within cellulosomes to close proximity and in proper orientation to the substrate for hydrolysis. In this dissertation, the family 10 xylanase from *C. thermocellum* incorporated into xylanosomes contained a family-22 CBM with known binding affinity to xylan [22]. However, a more strategic position of the CBM may be within the scaffolding protein instead of the enzyme for xylan hydrolysis. The location of the CBM can also be arranged at the N-terminal or C-terminal of the scaffolding, as well as in between cohesin domains, to place certain enzymes in better proximity to xylan. For instance, xylanases act on the insoluble xylan, whereas beta-xylosidases are active on soluble, smaller xylo-oligosaccharides. Furthermore, a comparison of xylanosomes with CBMs as a part of scaffolding proteins versus modular enzymes will provide further insight how to design xylanosomes for optimal substrate positioning, especially on substrates with varied xylan composition and side-chain substitution. Lastly, different xylan-binding CBMs, such as family-22 from *XynC* of *C. thermocellum* and family-35 from Fee1B of *Cellvibrio japonicus*, can be interchanged or combined, for assessing the most effective CBMs for xylan hydrolysis [23].

6.3.2 Development of a designer enzyme cocktail using cellulosomes and xylanosomes

Hemicellulases have been incorporated into tri-functional designer cellulosomes for enhanced synergy on biomass substrates, such as wheat straw [4]. With the knowledge gained from this dissertation, combinations of both mini-cellulosomes and xylanosomes can be formulated for enzymatic hydrolysis of biomass, therefore, taking advantage of the synergy of placing cellulases and hemicellulases in structured form. Based on results observed on destarched corn bran with xylanosomes, directed hydrolysis of hemicellulose with xylanosomes and cellulose with designer cellulosomes may result in a significant improvement in the release of fermentable sugars. Challenges would require development of chimeric scaffolding proteins with a number of cohesins from different cellulosome-producing species. A solution might require genome sequencing of more species or formation of designer xylanosomes and cellulosomes prior to addition to biomass for proper construction.

6.3.3 Engineering whole-cell biocatalysts for consolidated bioprocessing

Whole-cell bio-catalysis in biomass utilization can be defined engineering microbes to perform both saccharification and fermentation of biomass into desired products. Very recently, a xylanolytic binary culture of *E. coli* was developed that successfully expressed hemicellulases and fermented released sugars to ethanol [24]. Furthermore, mini-cellulosomes were constructed on the surface of yeast by expressing of chimeric scaffolding proteins onto the cell surface [7, 8]. There are many advantages in developing these types of whole-cell biocatalysts, including eliminating purification or concentration of enzyme cocktails, as well as combining biomass hydrolysis and conversion of sugars into a single bioprocessing step. Chapter 3 provided evidence for cell-surface display of xylanosomes through the expression of the two-cohesin

scaffolding protein, SP2 to the outer membrane of *E. coli*. Continued development to include extracellular expression of dockerin-tagged enzymes would result in creation of a xylanosome-producing *E. coli* strain that could directly hydrolyze xylan. This engineered strain could be combined with other commercially-available bacterial strains that ferment xylose to ethanol, such as *E. coli* K011 [25], or lactic acid, such as *E. coli* SZ63 [26]. Alternatively, the xylanosome-producing *E. coli* strain could be further engineered to effectively ferment xylose or other xylo-oligosaccharides for ethanol or other fermentation products. In conjunction with further development of the design and composition of xylanosomes, efficiency of engineered xylanolytic bacterial strains will continue to be enhanced.

Execution of the experiments outlined in this section will undoubtedly increase knowledge in xylan hydrolysis using xylanosomes. Furthermore, results will expand the arsenal of tools for the enzymatic hydrolysis of biomass and lead to more effective, and hopefully economical, consolidated bioprocessing strategies.

6.4 References

1. Arai, T., et al., *Synthesis of Clostridium cellulovorans minicellulosomes by intercellular complementation.* Proc Natl Acad Sci U S A, 2007. **104**(5): p. 1456-60.

2. Fierobe, H.P., et al., *Degradation of cellulose substrates by cellulosome chimeras. Substrate targeting versus proximity of enzyme components.* J Biol Chem, 2002. **277**(51): p. 49621-30.

3. Fierobe, H.P., et al., *Design and production of active cellulosome chimeras. Selective incorporation of dockerin-containing enzymes into defined functional complexes.* J Biol Chem, 2001. **276**(24): p. 21257-61.

4. Fierobe, H.P., et al., *Action of designer cellulosomes on homogeneous versus complex substrates: controlled incorporation of three distinct enzymes into a defined trifunctional scaffoldin.* J Biol Chem, 2005. **280**(16): p. 16325-34.

5. Mingardon, F., et al., *Exploration of new geometries in cellulosome-like chimeras.* Appl Environ Microbiol, 2007. **73**(22): p. 7138-49.

6. Mingardon, F., et al., *Heterologous production, assembly, and secretion of a minicellulosome by Clostridium acetobutylicum ATCC 824.* Appl Environ Microbiol, 2005. **71**(3): p. 1215-22.

7. Tsai, S.L., et al., *Functional assembly of minicellulosomes on the Saccharomyces cerevisiae cell surface for cellulose hydrolysis and ethanol production.* Appl Environ Microbiol, 2009. **75**(19): p. 6087-93.

8. Wen, F., J. Sun, and H. Zhao, *Yeast surface display of trifunctional minicellulosomes for simultaneous saccharification and fermentation of cellulose to ethanol.* Appl Environ Microbiol. **76**(4): p. 1251-60.

9. Schaeffer, F., et al., *Duplicated dockerin subdomains of Clostridium thermocellum endoglucanase CelD bind to a cohesin domain of the scaffolding protein CipA with distinct thermodynamic parameters and a negative cooperativity.* Biochemistry, 2002. **41**(7): p. 2106-14.

10. Carvalho, A.L., et al., *Evidence for a dual binding mode of dockerin modules to cohesins.* Proc Natl Acad Sci U S A, 2007. **104**(9): p. 3089-94.

11. DeBoy, R.T., et al., *Insights into plant cell wall degradation from the genome sequence of the soil bacterium Cellvibrio japonicus.* J Bacteriol, 2008. **190**(15): p. 5455-63.

12. Wamalwa, B.M., et al., *High-level heterologous expression of Bacillus halodurans putative xylanase xyn11a (BH0899) in Kluyveromyces lactis.* Biosci Biotechnol Biochem, 2007. **71**(3): p. 688-93.

13. Wagschal, K., et al., *Biochemical characterization of a novel dual-function arabinofuranosidase/xylosidase isolated from a compost starter mixture.* Appl Microbiol Biotechnol, 2009. **81**(5): p. 855-63.

14. Blum, D.L., et al., *Feruloyl esterase activity of the Clostridium thermocellum cellulosome can be attributed to previously unknown domains of XynY and XynZ.* J Bacteriol, 2000. **182**(5): p. 1346-51.

15. Kosugi, A., K. Murashima, and R.H. Doi, *Xylanase and acetyl xylan esterase activities of XynA, a key subunit of the Clostridium cellulovorans cellulosome for xylan degradation.* Appl Environ Microbiol, 2002. **68**(12): p. 6399-402.

16. Caspi, J., et al., *Effect of linker length and dockerin position on conversion of a Thermobifida fusca endoglucanase to the cellulosomal mode.* Appl Environ Microbiol, 2009. **75**(23): p. 7335-42.

17. Jiang, Z., et al., *Subunit composition of a large xylanolytic complex (xylanosome) from Streptomyces olivaceoviridis E-86.* J Biotechnol, 2006. **126**(3): p. 304-12.

18. Jiang, Z.Q., et al., *Characterization of a novel, ultra-large xylanolytic complex (xylanosome) from Streptomyces olivaceoviridis E-86.* Enzyme and Microbial Technology, 2005. **36**(7): p. 923-929.

19. Jiang, Z.-Q., et al., *A novel, ultra-large xylanolytic complex (xylanosome) secreted by <i>Streptomyces olivaceoviridis</i&gt.* Biotechnology Letters, 2004. **26**(5): p. 431-436.

20. Raman, B., et al., *Impact of pretreated Switchgrass and biomass carbohydrates on Clostridium thermocellum ATCC 27405 cellulosome composition: a quantitative proteomic analysis.* PLoS One, 2009. **4**(4): p. e5271.

21. Shallom, D. and Y. Shoham, *Microbial hemicellulases.* Curr Opin Microbiol, 2003. **6**(3): p. 219-28.

22. Ali, E., et al., *Functions of family-22 carbohydrate-binding modules in Clostridium josui Xyn10A.* Biosci Biotechnol Biochem, 2005. **69**(12): p. 2389-94.

23. McCartney, L., et al., *Differential recognition of plant cell walls by microbial xylan-specific carbohydrate-binding modules.* Proc Natl Acad Sci U S A, 2006. **103**(12): p. 4765-70.

24. Shin, H.D., et al., *Escherichia coli binary culture engineered for direct fermentation of hemicellulose to a biofuel.* Appl Environ Microbiol. **76**(24): p. 8150-9.

25. Ohta, K., et al., *Genetic improvement of Escherichia coli for ethanol production: chromosomal integration of Zymomonas mobilis genes encoding pyruvate decarboxylase and alcohol dehydrogenase II.* Appl Environ Microbiol, 1991. **57**(4): p. 893-900.

26. Zhou, S., K.T. Shanmugam, and L.O. Ingram, *Functional replacement of the Escherichia coli D-(-)-lactate dehydrogenase gene (ldhA) with the L-(+)-lactate dehydrogenase gene (ldhL) from Pediococcus acidilactici.* Appl Environ Microbiol, 2003. **69**(4): p. 2237-44.